E&M ECONOMICS AND MANAGEMENT

E+M EKONOMIE A MANAGEMENT

Economics
Business Administration & Management
Marketing & Trade
Finance
Information Management

volume XXIX
2026
2

Cooperating faculties

Faculty of Economics
Technical University of Liberec

FACULTY OF ECONOMICS TUL

Liberec
Czech Republic
https://www.ef.tul.cz/

Faculty of Economics and Administration
University of Pardubice

UNIVERSITY OF PARDUBICE
FACULTY OF ECONOMICS AND ADMINISTRATION

Pardubice
Czech Republic
https://fes.upce.cz

Faculty of Economics
University of West Bohemia in Pilsen

FACULTY OF ECONOMICS
UNIVERSITY OF WEST BOHEMIA

Plzeň
Czech Republic
https://www.fek.zcu.cz

Faculty of Management and Economics
Tomas Bata University in Zlín

Tomas Bata University
Faculty of Management and Economics

Zlín
Czech Republic
https://fame.utb.cz

Faculty of Informatics and Management
University of Hradec Králové

University of Hradec Králové
Faculty of Informatics and Management

Hradec Králové
Czech Republic
https://www.uhk.cz/fim

Faculty of Social and Economic Studies
Jan Evangelista Purkyně University in Ústí nad Labem

JAN EVANGELISTA PURKYNĚ UNIVERSITY IN ÚSTÍ NAD LABEM
Faculty of Social and Economic Studies

Ústí nad Labem
Czech Republic
https://www.fse.ujep.cz

School of Business Administration in Karviná
Silesian University in Opava

SILESIAN UNIVERSITY
SCHOOL OF BUSINESS ADMINISTRATION IN KARVINA

Karviná
Czech Republic
https://www.slu.cz/opf/cz

Faculty of Economics
Technical University of Košice

TECHNICAL UNIVERSITY OF KOŠICE
Faculty of Economics

Košice
Slovak Republic
https://ekf.tuke.sk

Faculty of Economics
Matej Bel University in Banská Bystrica

FACULTY OF ECONOMICS
MATEJ BEL UNIVERSITY IN BANSKÁ BYSTRICA

Banská Bystrica
Slovak Republic
https://www.ef.umb.sk

Ekonomie a Management
Economics & Management

scientific journal
www.ekonomie-management.cz

The journal E&M Economics and Management (E&M) publishes high quality original research articles and scientific studies based on theoretical and empirical analyses. E&M encourages new intriguing ideas and new perspectives on existing state of knowledge. Among the key topics covered are Economics, Business Administration, Finance, Management, Information Management, and Marketing & Trade. The journal features theoretical articles as well as application-oriented papers. Occasionally published review articles summarising existing knowledge are accepted only if they are based on a systematic literature review. The journal is published quarterly and all articles are double-blind peer reviewed. Editorial board consistently pays attention to compliance with the rules of publishing ethics. Therefore, among other things, articles received are checked by the Crossref's service Similarity Check.

The journal is covered in the Social Sciences Citation Index, Social SciSearch and Journal Citation Reports/ /Social Sciences Edition (https://clarivate.com). It is also monitored by Google Scholar (https://scholar.google.com/), Research Papers in Economics (RePEc; https://econpapers.repec.org/article/), the electronic EconLit index (https://www.aeaweb.org/econlit/), International Bibliography of the Social Sciences (http://www.proquest.com/products-services/ibss-set-c.html) and by SCOPUS (https://www.elsevier.com/solutions/scopus), ABI/INFORM (http://www.proquest.com/products-services/abi_inform_complete.html), EBSCO Information Services (https://www.ebsco.com) and GALE (https://support.gale.com/tlist/products) databases. It is listed in the 11th Edition of Cabell's Directory of Publishing Opportunities in Economics and Finance/Management (www.cabells.com).

Contents

Economics

Business administration and management

Environmental innovation and green low-carbon transitions: The moderating role of green regulatory pressure

***Tingqian Pu*[1]**

[1] *Guizhou University of Commerce, School of Economics and Finance, Department of Financial Engineering, China, ORCID: 0000-0002-6618-5891, tq_pu@outlook.com.*

Abstract: *Environmental innovation is widely recognized as a key driver of sustainability, reducing greenhouse gas emissions and enhancing resource efficiency. However, the mechanisms through which it contributes to green low-carbon transitions remain underexplored, particularly in the context of varying regulatory environments. This study examines the relationship between environmental innovation and green low-carbon transitions, focusing on the moderating role of green regulatory pressure. Using panel data from 277 prefecture-level cities in China between 2012 and 2021, the analysis reveals that environmental innovation significantly advances green low-carbon transitions, demonstrating its transformative potential in driving sustainable economic development. However, green regulatory pressure negatively moderates this relationship, diminishing the positive effects of environmental innovation. These findings underscore the dual-edged nature of green regulatory pressure, highlighting how excessive regulatory intensity can constrain innovation effectiveness. By providing a comprehensive analysis of the interaction between environmental innovation and regulation, this study offers actionable insights for optimizing environmental policies and addresses critical research gaps, particularly in the context of emerging economies.*

Keywords: *Environmental innovation, green low-carbon transition, green regulatory pressure, environmental policies, emerging economies.*

JEL Classification: *O32, O33, Q56, Q58.*

APA Style Citation: Pu, T. (2026). Environmental innovation and green low-carbon transitions: The moderating role of green regulatory pressure. *E&M Economics and Management*, *29*(2), 1–15. https://doi.org/10.15240/tul/001/2025-5-021

Early Access Publication Date: *August 1, 2025.*

Introduction

The transition to a green low-carbon economy has become a global imperative in response to mounting environmental challenges and the urgency of mitigating climate change. Central to this transition is the adoption of green low-carbon transitions (GLT), which requires systemic changes in industrial practices, energy use, and production processes (Wang et al., 2017). Environmental innovation (EI), encompassing technological advancements and process improvements aimed at reducing environmental harm, has been increasingly recognized as a critical mechanism in driving this transformation (Xu, 2024). By reducing greenhouse gas emissions, enhancing energy efficiency, and fostering cleaner production processes, EI enables firms to align with sustainability goals while maintaining competitive advantage (Anu et al., 2022). However, despite extensive acknowledgment of EI's pivotal role in achieving green and sustainable development, the mechanisms through which it contributes to broader low-carbon transitions remain underexplored, leaving notable gaps in understanding its systemic impact.

Building on EI's recognized benefits, existing research has primarily examined its role in specific contexts, highlighting its diverse contributions to sustainability. For example, studies have documented the role of EI in reducing carbon emissions (Albitar et al., 2022), improving resource efficiency (Aftab et al., 2022), and fostering cleaner production practices (Mora-Contreras et al., 2023). Additionally, broader investigations into technological innovation have emphasized its transformative role in facilitating systemic shifts toward low-carbon economies (Yang et al., 2023). While these studies underscore the environmental and operational advantages of EI, they often focus on immediate or localized outcomes rather than its long-term impact on GLT. Furthermore, much of this research is concentrated in developed economies, with limited exploration of emerging markets like China, where unique challenges such as rapid industrialization and diverse regulatory frameworks offer opportunities for deeper insights. This gap underscores the need for more nuanced investigations into how EI drives GLT across varying economic and regulatory contexts.

To address these gaps, this study examines how EI influences GLT and explores the moderating role of green regulatory pressure (GRP), a critical external factor shaping the regulatory environment for innovation. GRP has the potential to enhance sustainability outcomes by incentivizing firms through compliance mechanisms, but it may also impose costs and constraints that limit the effective implementation of innovative solutions (Xu et al., 2022; Yu et al., 2023). China presents an ideal context for this analysis due to its rapid industrialization, ambitious environmental goals, and diverse regulatory landscapes. As the largest developing economy and the world's leading emitter of greenhouse gases, China faces the dual challenge of sustaining economic growth while addressing environmental degradation. Its dual-carbon targets (peaking carbon emissions by 2030 and achieving carbon neutrality by 2060) demonstrate an unwavering commitment to green transformation. Furthermore, China's decentralized governance structure creates significant regional variations in regulatory intensity, offering a natural case to examine how GRP moderates the EI-GLT relationship.

Using panel data from 277 prefecture-level cities in China over the period 2012–2021, this study investigates the relationship between EI and GLT. The findings reveal that EI significantly promotes GLT, underscoring its transformative role in advancing low-carbon transitions. However, GRP negatively moderates this relationship, reducing the positive impact of EI on GLT. These results provide nuanced insights into how environmental technological innovation and regulation interact in shaping sustainable development and highlight the importance of contextual factors in determining policy effectiveness.

This study contributes to the literature and policy discussions in several ways. First, it expands the theoretical understanding of how EI drives GLT by elucidating the mechanisms linking innovation to sustainable development. Second, the research integrates EI and GRP into a unified analytical framework, offering a comprehensive view of how external regulatory environments influence the effectiveness of innovation in sustainability outcomes. Third, the findings provide actionable policy recommendations for balancing GRP intensity to maximize its positive effects on EI while minimizing unintended constraints, thereby informing the design of more effective environmental policies.

The remainder of this paper is organized as follows. Section 1 outlines the theoretical underpinnings and develops the research hypotheses. Section 2 details the data sources and empirical methodology used to analyze the relationships among the key variables. Section 3 presents the empirical results and robustness checks. The last section concludes by summarizing the key findings, discussing their implications, and proposing policy recommendations and avenues for future research.

1 Theoretical foundations and hypotheses

1.1 Environmental innovation and green low-carbon transition

The transition toward green low-carbon economies has increasingly been driven by the strategic adoption of EI. From the perspective of resource-based theory (RBT), firms' ability to navigate this transition is rooted in the development and deployment of unique, valuable, and inimitable resources that align with sustainability goals (Gargallo et al., 2024). EI, which encompasses technological advancements, cleaner production processes, and energy-efficient technologies, constitutes a critical resource for achieving a green low-carbon transition (Jie et al., 2023). By integrating these innovations, firms can simultaneously reduce environmental harm and enhance operational efficiency, establishing a competitive edge in increasingly eco-conscious markets.

While RBT highlights the intrinsic value of EI as a firm-level resource, innovation adoption frameworks emphasize the dynamic role of innovation in enabling systemic transformations (Cannavacciuolo et al., 2023; Pu, 2025). Zhao et al. (2024) posit that EI facilitates a "double dividend," achieving dual objectives of environmental sustainability and economic efficiency. This perspective aligns with Schumpeterian views of creative destruction, where innovation serves as a mechanism for disrupting entrenched practices and paving the way for more sustainable production systems (e.g., Khan et al., 2023; Lile et al., 2024). Thus, EI is not only firm-specific resources but also transformative agents that reshape industry practices, reduce carbon emissions, and accelerate the global shift toward low-carbon economies.

These theoretical perspectives converge to underscore the pivotal role of EI in driving GLT. On the one hand, RBT positions EI as a resource that enables firms to leverage environmental capabilities for competitive advantage. On the other hand, innovation adoption theories highlight its systemic impact, demonstrating that widespread adoption of green innovations can catalyze broader economic and environmental transitions. Empirical studies provide robust support for these claims. Xu et al. (2021) find that firms investing in EI achieve substantial reductions in greenhouse gas emissions while enhancing resource efficiency. Similarly, Li et al. (2024) demonstrate that eco-innovations contribute significantly to firms' environmental and operational performance, positioning them as leaders in the low-carbon transition.

While RBT focuses on the development of internal capabilities that enable firms to pursue sustainability-oriented strategies, it does not fully account for the external institutional contexts that shape how these capabilities are activated and leveraged. In contrast, institutional theory (IT) emphasizes the role of formal rules, regulatory structures, and normative expectations in influencing organizational behavior, particularly under conditions of environmental scrutiny and societal pressure. By integrating these perspectives, we view EI as a firm-specific strategic resource (as per RBT), whose effectiveness is contingent upon the external regulatory environment (as per IT). This integration allows for a more comprehensive understanding of how internal resources interact with external constraints to influence GLT. It thus provides a robust theoretical foundation for expecting a positive relationship between EI and GLT, as formally stated in the following hypothesis:

H1: EI has a positive and significant impact on GLT.

1.2 The moderating role of green regulatory pressure

While EI serves as a catalyst for GLT, the role of external regulatory forces (specifically GRP) introduces complexities that can either amplify or dampen this relationship. IT provides a useful lens to understand this dynamic. Coercive institutional pressures, such as stringent environmental regulations, compel firms to adopt sustainable practices to comply with legal and societal expectations (Burdon & Sorour, 2018). However, the effects of such pressures are contingent upon their design, intensity, and the regulatory environment's overall flexibility.

The "Porter hypothesis" offers nuanced insights into the role of environmental regulations. Porter and Linde (1995) argue that well-designed regulations can stimulate innovation and improve firm competitiveness by encouraging the development of resource-efficient technologies. However, this hypothesis also acknowledges that excessive or poorly structured regulations may lead to compliance burdens that inhibit innovation (Ambec et al., 2013). For example, stringent regulatory requirements may divert resources away from innovation activities, while frequent policy changes can create uncertainty, deterring firms from long-term investments in EI (York & Venkataraman, 2010).

Empirical studies provide mixed evidence on the moderating effects of GRP. Wang et al. (2022) show that moderate regulatory pressure enhances firms' green innovation performance, yet excessive pressure can lead to diminishing returns by constraining firms' operational flexibility and increasing compliance costs.

Similarly, Forés (2024) finds that while GRP incentivizes firms to adopt green practices, it may negatively interact with innovation efforts in resource-intensive industries, where compliance costs are particularly high.

The shift in GRP's role (from enabling to constraining) is dependent on a range of contextual factors. When regulatory frameworks are transparent, stable, and aligned with firms' strategic objectives, they can enhance innovation by reducing uncertainty and legitimizing green investment. Conversely, when regulations are overly rigid, inconsistently applied, or lack supporting mechanisms such as financial incentives or technical guidance, they may produce unintended consequences. Under such conditions, firms may redirect resources toward short-term compliance rather than long-term innovation, especially if they operate in highly regulated or capital-intensive industries. This contingency reflects the expanded interpretation of the Porter hypothesis and aligns with IT's recognition that institutional pressures may have both facilitating and inhibiting effects.

In the context of GLT, GRP is likely to act as a double-edged sword. While regulatory frameworks are essential for creating incentives for EI, overly stringent regulations may reduce firms' ability to translate innovation into measurable green outcomes. This occurs because excessive regulatory intensity can shift managerial focus toward meeting compliance requirements rather than implementing and optimizing innovative solutions. Thus, we propose the following hypothesis:

H2: GRP moderates the relationship between EI and GLT.

2 Research methodology

2.1 Data source and processing

The data used in this study spans the period from 2012 to 2021 and focuses on 277 prefecture-level cities in China. The sample focuses on prefecture-level cities because they serve as critical hubs of population density, industrial activity, and environmental governance. These cities play a pivotal role in balancing economic growth with environmental sustainability, making them ideal units for evaluating the effectiveness of EI and the moderating role of GRP in promoting GLT. The diversity in regulatory intensity and industrial composition across cities offers a robust empirical setting to examine the interplay between innovation and regulation.

The dataset integrates information from multiple authoritative sources, including the China Urban Statistical Yearbook, the China Provincial Statistical Yearbook, and regional government work reports. The statistical yearbooks provide detailed data on socio-economic indicators, energy consumption, and industrial output. Local government work reports, which offer qualitative insights into environmental policies and governance strategies, were manually collected from official municipal websites and online archives. Regions with missing or inconsistent data across key variables were excluded from the final sample to ensure the integrity and reliability of the analysis.

2.2 Variables construction

(I) Dependent variable. Following the research of Feng and Ge (2024), *GLT* is conceptualized in this study as a multidimensional framework encompassing four core aspects: carbon emission reduction (CER), pollutants emission reduction (PER), energy saving efficiency (ESE), and transformation performance as measured by green low-carbon total factor productivity (GTFP). The first three dimensions (CER, PER, and ESE) reflect the effectiveness of green low-carbon initiatives in mitigating environmental impacts, while GTFP evaluates the overall transformation benefits. To ensure consistency in measurement CO_2 emissions and energy emissions have been logarithmized. Comprehensive pollutant emissions are quantified by aggregating industrial wastewater, industrial soot, and sulfur dioxide (SO_2) emissions using the entropy weighting method, thereby providing a more balanced representation of various pollutants.

Additionally, we employ an advanced data envelopment analysis-slack based measure (DEA-SBM) model, combined with the global Malmquist-Luenberger (GML) index, to quantify *GLT* across Chinese cities. The DEA-SBM model incorporates the non-angular, non-radial directional distance function (DDF) proposed by Zhou et al. (2012), enabling a more nuanced evaluation of urban efficiency by explicitly accounting for both desired and undesired outputs. Desired outputs include gross domestic product (GDP), whereas undesired outputs encompass emissions such as industrial wastewater, industrial soot, SO_2, and CO_2. CO_2 emissions are estimated based on energy consumption data obtained from the China Energy Statistical

Yearbook, which includes natural gas, liquefied petroleum gas, electricity, and heat energy. These data are combined with officially announced emission factors to estimate carbon emissions at the city level (Wu et al., 2016). Input indicators for *GLT* consist of capital, labor, and energy. Labor input is quantified as the number of employees at the prefecture level, while energy input is represented by total electricity consumption. By integrating these input factors, desired outputs, and undesired outputs, the DEA-SBM model, in conjunction with the GML index, provides a comprehensive metric for assessing *GLT*.

(II) Independent variable. The primary independent variable in this study is *EI*. *EI* refers to the development and application of new or significantly improved products, processes, and systems aimed at minimizing environmental impacts and fostering sustainable development. Unlike traditional technological innovation, which often neglects ecological consequences, *EI* directly targets sustainability through advancements in resource efficiency, pollution prevention, and emission reduction, playing a crucial role in the transition to a green, low-carbon economy (Owen et al., 2018). This concept includes innovations in renewable energy, energy efficiency, pollution control, and sustainable production methods, all of which contribute significantly to environmental improvement. Following the method of Többelmann and Wendler (2019), this study quantifies *EI* by using the number of green patents issued per 10,000 people, obtained from the CNRDS Database. Green patents are a key indicator of environmentally directed technological progress, capturing the scale and impact of innovation efforts in achieving sustainability goals.

(III) Moderating variable. The moderating variable in this study is *GRP*. *GRP* reflects the extent to which government interventions (including policy regulations, environmental standards, and incentive mechanisms) guide and constrain economic actors towards sustainable development goals. Unlike traditional conceptualizations of government green attention, *GRP* emphasizes not only the government's focus on ecological issues but also the pathways through which regulatory mechanisms are enforced and transmitted. This regulatory dimension encapsulates the diverse tools used by the government to ensure adherence to green practices across multiple administrative levels, representing a crucial moderator of green technological innovation and its subsequent impact on low-carbon transitions.

To quantitatively assess *GRP*, this study employed a text-mining approach on government policy documents and annual work reports from 2012 to 2021 across 277 prefecture-level cities in China. By analyzing policy texts, we extracted and categorized key terms representing green regulatory initiatives under four thematic dimensions: environmental regulation mechanism, green governance, low-carbon economy, and resource utilization and ecological protection.

Utilizing a neural network-based Word2Vec model, we expanded the initial seed list of green-related keywords to include semantically related terms, thereby ensuring the comprehensiveness of the green regulatory lexicon. Python tools such as Jieba and Gensim were employed to tokenize the texts and build the language models. Both continuous bag-of-words (CBOW) and skip-gram methods were applied to identify hidden relationships among terms and enrich the vocabulary beyond initial boundaries.

For each city-year, the total frequency of green regulation-related keywords was calculated across all policy documents. These frequencies were then log-transformed using the formula ln (keyword frequency + 1) to construct a standardized *GRP* index. All keywords were assigned equal weight in constructing the *GRP* index. This decision was made to avoid introducing subjective judgments regarding the relative importance or regulatory strength of specific terms, and to ensure consistency and comparability across cities and over time. We acknowledge that certain expressions (such as those indicating quantifiable policy targets or binding financial commitments) may reflect stronger regulatory intent than general recommendations. However, the purpose of the *GRP* variable is to capture the overall salience of green regulatory discourse rather than assess the enforceability of individual statements. A detailed breakdown of the keyword categories and their thematic classifications is provided in Tab. 1.

(IV) Control variables. Following prior studies (e.g., Baran et al., 2020; Feng & Ge, 2024; Luo et al., 2022), several control variables are incorporated into the analysis to account for

Tab. 1: Keywords of green regulatory pressure

Dimension	Keywords
Environmental regulation mechanism	Environmental regulation, green governance, energy conservation and emission reduction, industrial water saving, emission standards, regulatory policies, carbon trading mechanism, compliance enforcement
Green governance	Green governance, green economy, ecological civilization, eco-city, circular economy, green manufacturing, policy incentives, regulatory framework, ecological audits
Low-carbon economy	Low-carbon economy, energy reduction, high energy consumption, green innovation, renewable energy targets, emission reduction targets, decarbonization policies
Resource utilization and ecological protection	Water-saving irrigation, agricultural non-point source pollution control, forest restoration, water source conservation, afforestation, waste management, green mountains and rivers, forest preservation, resource conservation policies

Source: own

potential confounding factors that may influence the relationship between *EI* and *GLT*. Financial support (*FS*) is measured as the ratio of loans disbursed by financial institutions to year-end GDP, reflecting the availability of credit and overall financial capacity within the region. This variable captures the extent to which financial resources can drive investment in innovation and low-carbon projects. Mining employment share (*MES*), defined as the percentage of total employees in the mining industry, serves as a proxy for the region's reliance on resource-intensive economic activities. This variable reflects the structural characteristics of the economy, which may affect the feasibility and speed of transitioning toward a low-carbon model. Urbanization level (*UL*) is measured as the ratio of the urban population to the total population, representing the degree of urbanization within a region. Higher levels of urbanization are often associated with increased resource efficiency and innovation diffusion, which can influence the effectiveness of environmental policies. Government intervention (*GI*), expressed as the ratio of fiscal spending to GDP, provides a measure of the government's involvement in economic activities. Greater fiscal spending can indicate proactive public-sector support for low-carbon initiatives, but it may also introduce distortions that affect market-driven innovation. Population density (*PD*), calculated as the proportion of the regional population to the total land area, reflects the intensity of human activity within a region. Higher population density is often linked to greater environmental pressures but can also enhance the diffusion of innovative practices due to economies of scale and network effects. These control variables offer a comprehensive framework to account for external influences that could confound the empirical analysis, ensuring robust estimation of the relationship between *EI*, *GRP*, and *GLT*.

2.3 Model specification

To examine the impact of *EI* on *GLT* and the moderating role of *GRP*, we constructed the following panel data fixed-effects ordinary least squares models:

$$GLT_{i,t} = \alpha_0 + \alpha_1 EI_{i,t} + \alpha_2 Controls_{i,t} + \\ + \alpha_3 Year_{i,t} + \alpha_4 City_{i,t} + \varepsilon \tag{1}$$

$$GLT_{i,t} = \alpha_0 + \alpha_1 EI_{i,t} + \alpha_2 GRP_{i,t} + \\ + \alpha_3 EI_{i,t} * GRP_{i,t} + \alpha_4 Controls_{i,t} + \\ + \alpha_5 Year_{i,t} + \alpha_6 City_{i,t} + \varepsilon \tag{2}$$

where: α_0 – the intercept, while $\alpha_1 - \alpha_6$ – the coefficients to be estimated; the term ε – the error term; the subscript *i* – the cross-sectional dimension of cities; *t* – the time-series dimension. The interaction term $EI_{i,t} * GRP_{i,t}$ captures the moderating effect of green regulatory pressure on the relationship between environmental innovation and green low-carbon transition. Controls include a set of control variables as previously described, accounting for potential confounding factors that could influence *GLT*.

Furthermore, we incorporated city-specific fixed effects to account for unobservable heterogeneity across cities, such as geographic or cultural factors that are constant over time but may affect the *GLT*. Similarly, year fixed effects were included to control for macroeconomic trends, policy changes, or external shocks that could simultaneously affect all cities, ensuring that the temporal variation in *GLT* is not spuriously attributed to *EI* or *GRP*.

3 Empirical results

3.1 Descriptive statistics

Tab. 2 presents the descriptive statistics for the key variables used in the analysis, providing an overview of their distribution and variability across the sample of 2,770 observations.

The dependent variable, *GLT*, exhibits a mean value of 0.855 with a standard deviation of 0.261, indicating moderate variability among regions. The range of *GLT* spans from 0.320 to 1.683, reflecting substantial differences in the extent of low-carbon transitions across prefecture-level cities.

The independent variable, *EI*, has an average value of 0.108 with a standard deviation of 0.187, suggesting considerable heterogeneity in the adoption of environmentally innovative practices. The minimum value is 0.000, while the maximum reaches 0.922, highlighting disparities in *EI* intensity across cities.

The moderating variable, *GRP*, displays a mean of 6.056 and a standard deviation of 3.558. The range from 0.000 to 16.000 indicates substantial variation in the intensity of regulatory efforts among regions, capturing diverse regulatory environments.

Among the control variables, *FS* has a mean of 0.941, indicating variations in the availability of credit, with a range of 0.368 to 2.702. *MES* averages 0.049, reflecting regional dependence on resource-based industries, while *UL* shows significant diversity, with a mean of 53.657 and values ranging from 27.350 to 88.820. *GI*, measured as fiscal spending to GDP, averages 0.182, highlighting differences in public sector involvement, and *PD*, with a mean of 0.045, captures demographic disparities across regions. These variables provide critical context for understanding the economic, industrial, and demographic factors influencing green low-carbon transition.

3.2 Correlation and variance inflation factor analysis

Tab. 3 reports the results of the Pearson correlation analysis, highlighting significant relationships among the key variables of interest. *EI* shows a positive and significant correlation with *GLT* ($r = 0.214$, $p < 0.01$), indicating that higher levels of *EI* are associated with greater advancements in low-carbon transitions. Conversely, *GRP* exhibits a weak negative correlation with *GLT* ($r = -0.055$, $p < 0.01$), suggesting that increased regulatory pressure may slightly hinder the progress of low-carbon transitions.

Tab. 2: Descriptive statistics of main variables

Variable	Obs	Mean	Std. dev.	Min	Max
GLT	2,770	0.855	0.261	0.320	1.683
EI	2,770	0.108	0.187	0.000	0.922
GRP	2,770	6.056	3.558	0.000	16.000
FS	2,770	0.941	0.508	0.368	2.702
MES	2,770	0.049	0.081	0.000	0.333
UL	2,770	53.657	14.418	27.350	88.820
GI	2,770	0.182	0.090	0.035	0.458
PD	2,770	0.045	0.038	0.003	0.196

Note: *GLT* – green low-carbon transition; *EI* – environmental innovation; *GRP* – green regulatory pressure; *FS* – financial support; *MES* – mining employment share; *UL* – urbanization level; *GI* – government intervention; *PD* – population density.

Source: own

Tab. 3: Pearson correlation and variance inflation factor analysis

	GLT	*EI*	*GRP*	*FS*	*MES*	*UL*	*GI*	*PD*	VIF
GLT	1.000								
EI	0.214***	1.000							2.15
GRP	−0.055***	−0.035*	1.000						1.03
FS	0.150***	0.519***	−0.072***	1.000					1.67
MES	0.014	−0.199***	−0.037*	−0.146***	1.000				1.09
UL	0.149***	0.601***	−0.064***	0.455***	−0.048**	1.000			1.90
GI	−0.148***	−0.214***	−0.117***	0.149***	0.019	−0.256***	1.000		1.34
PD	0.154***	0.554***	0.007	0.208***	−0.203***	0.515***	−0.381***	1.000	1.77

Note: * $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Source: own

However, the ultimate nature of these relationships requires further validation through regression analysis.

The variance inflation factor (VIF) values for *EI* (2.15) and *GRP* (1.03) are well below the commonly accepted threshold of 10, indicating no multicollinearity concerns between these variables. These results validate the inclusion of *EI* and *GRP* in the regression models and provide a clear foundation for exploring their respective impacts on *GLT* in the subsequent analysis.

3.3 Univariate analysis

Tab. 4 presents the univariate analysis comparing *GLT* between cities with and without significant *EI*. Cities with *EI* (dummy *EI* = 1) exhibit a higher mean *GLT* value of 0.860, compared to 0.760 for cities without *EI* (dummy *EI* = 0). The difference in means is statistically significant ($t = 0.070$, $p < 0.01$), indicating that cities engaging in *EI* demonstrate greater advancements in *GLT*. Additionally, the positive sloping linear regression line in Fig. 1 indicates a positive correlation between *EI* and *GLT*. These findings suggest that cities undergoing *EI* exhibit higher levels of *GLT* compared to cities that have not engaged in *EI*.

Tab. 4: Univariate analysis

Variable	Dummy (*EI*) = 1		Dummy (*EI*) = 0		Differences *T*-value
	N	Mean	*N*	Mean	
GLT	2,574	0.860	196	0.790	0.070***

Note: *** denotes statistical significance at the 1% level.

Source: own

3.4 Multivariate results

Tab. 5 reports the baseline regression results examining the relationship between *EI* and *GLT*. The models progressively incorporate fixed effects and control variables to provide a comprehensive analysis of the factors influencing *GLT*. Robust standard errors are clustered at the city level to account for heteroskedasticity and within-group correlation.

Across all specifications in Tab. 5, *EI* consistently demonstrates a positive and significant impact on *GLT*, highlighting its critical role as a driver of sustainability. In Column (1), the simplest specification with *EI* as the sole independent variable, the coefficient for *EI* is positive and highly significant ($\alpha_1 = 0.300$, $p < 0.01$), indicating a strong direct association between *EI* and *GLT*. When year and city

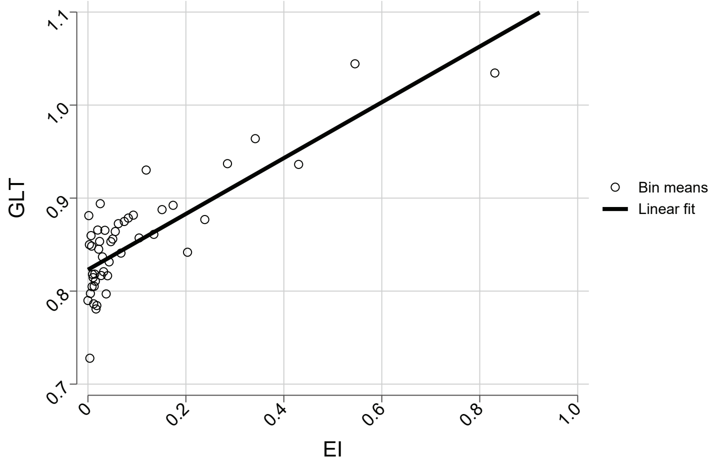

Fig. 1: The linear fit of *EI* and *GLT*

Note: Due to the large number of observations involved in the linear fitting of the sample, this study divides *GLT* into 50 equally sized groups for analysis in Fig. 1.

Source: own

fixed effects are added in Column (2), the coefficient for *EI* remains positive and significant (α_1 = 0.153, $p < 0.05$), though its magnitude decreases, suggesting that part of the observed variation is attributable to broader temporal and spatial factors. In Column (3), which introduces control variables such as *FS*, *MES*, *UL*, *GI*, and *PD*, the coefficient for *EI* remains robustly significant (α_1 = 0.202, $p < 0.01$), underscoring the persistent impact of *EI* on *GLT* even after accounting for potential confounders. Finally, in the fully specified model in Column (4), which includes both fixed effects and control variables, the coefficient for *EI* remains positive and significant (α_1 = 0.147, $p < 0.01$), reinforcing the robustness of its effect. These results provide strong evidence that *EI* plays a transformative role in advancing *GLT*, even under varying model specifications and after controlling for economic, industrial, and governance-related factors.

Overall, the findings strongly affirm the pivotal role of *EI* in driving green *GLT*, aligning with prior studies that emphasize its effectiveness in reducing carbon emissions, enhancing resource efficiency, and fostering sustainability (e.g., Khan et al., 2023; Wang et al., 2022; Wu et al., 2016). These studies underscore how technological advancements and process improvements enable firms to align economic objectives with environmental imperatives, thus paving pathways for sustainable economic growth. Consistent with this body of work, our results highlight the transformative potential of *EI*, demonstrating its capacity to facilitate systemic shifts in industrial practices, energy utilization, and production processes that are critical for achieving low-carbon objectives. However, what differentiates this analysis from existing literature is its contextual focus on emerging markets, particularly China's distinctive industrial and regulatory environment. Most previous research, concentrated in developed economies, emphasizes the direct environmental benefits of innovation while often overlooking the unique challenges and opportunities arising in rapidly industrializing economies (e.g., Cannavacciuolo et al., 2023; Ghisellini et al., 2015; Pu & Zulkafli, 2024). China, as the world's largest developing economy, presents an intricate balance of ambitious sustainability goals and diverse regional dynamics, offering

Tab. 5: Main results

Variable	(1)	(2)	(3)	(4)
	GLT	*GLT*	*GLT*	*GLT*
EI	0.300***	0.153***	0.202***	0.147***
	(0.026)	(0.039)	(0.040)	(0.041)
FS			−0.001	−0.004**
			(0.001)	(0.002)
MES			−0.384***	0.278**
			(0.070)	(0.115)
UL			0.059***	0.055***
			(0.015)	(0.019)
GI			0.223	0.553
			(0.164)	(1.501)
PD			0.216***	0.286*
			(0.061)	(0.162)
_cons	0.823***	0.839***	0.864***	0.892***
	(0.006)	(0.005)	(0.025)	(0.110)
Year FE	No	Yes	No	Yes
City FE	No	Yes	No	Yes
N	2,770	2,770	2,770	2,770
Adj. R^2	0.046	0.683	0.066	0.688

Note: Statistical significance is indicated by *, **, and *** for the 10, 5, and 1% levels, respectively; robustness standard errors are provided in brackets.

Source: own

a fertile ground for examining how *EI* operates under varying regulatory and economic conditions. This contextual nuance resonates with RBT and IT, providing a dual lens to understand how *EI* functions both as a strategic resource that enhances firms' competitive advantage and as a mechanism influenced by external institutional pressures.

3.5 Endogeneity test

To address potential endogeneity concerns and ensure the robustness of our findings, we employ four complementary approaches: the Driscoll-Kraay fixed effects (DK-FE) technique, a two-step system GMM estimator, a lagged dependent variable (DV) approach, and quantile regressions. These methods are designed to mitigate potential biases arising from simultaneity, omitted variables, and cross-sectional dependence.

(I) Driscoll-Kraay fixed effects (DK-FE). The DK-FE method adjusts for heteroscedasticity, autocorrelation, and cross-sectional dependence, issues commonly encountered in panel data with large cross-sectional dimensions. As shown in Column (1) of Tab. 6, the estimated coefficient for *EI* remains positive and statistically significant (coefficient = 0.227, p < 0.01), confirming the robustness of the baseline results under this specification.

(II) Two-step system generalized method of moments (GMM). To address simultaneity bias and unobserved firm-level heterogeneity, we utilize the two-step system GMM estimator. This method leverages lagged variables as instruments, ensuring consistent and efficient parameter estimation. Column (2) indicates that *EI* retains a significant and positive effect on *GLT* (coefficient = 0.261, p < 0.05). The Hansen J-statistic further validates the appropriateness

Tab. 6: Endogeneity test

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	DK-FE	GMM	Lagged DV	Quantile regressions				
	GLT	*GLT*	$GLT_{(t+1)}$	*GLT*	*GLT*	*GLT*	*GLT*	*GLT*
EI	0.227***	0.261**	0.098**	0.155**	0.236***	0.266***	0.255***	0.215***
	(0.019)	(0.129)	(0.047)	(0.072)	(0.041)	(0.029)	(0.028)	(0.030)
L.*GLT*		1.020***						
		(0.057)						
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AR (1)		0.658						
AR (2)		0.117						
Hansen-J		Yes						
N	2,770	2,493	2,493	2,770	2,770	2,770	2,770	2,770
Adj. R^2	0.202		0.752	0.111	0.121	0.130	0.136	0.181

Note: Statistical significance is indicated by *, **, and *** for the 10, 5, and 1% levels, respectively; robustness standard errors are provided in brackets.

Source: own

of the instrumental variables, alleviating concerns regarding model specification.

(III) Lagged dependent variable (DV). Incorporating a lagged dependent variable helps capture dynamic relationships and addresses potential reverse causality between *EI* and *GLT*. Although L.*GLT* appears in the system GMM model (Column 2 of Tab. 6), its inclusion provides evidence of strong persistence in *GLT* over time. Importantly, in the separate model using *GLT* at time $t + 1$ as the dependent variable (Column 3 of Tab. 6), *EI* remains positively and significantly associated with future *GLT* levels (coefficient = 0.098, $p < 0.05$), further supporting the robustness of the main findings.

(IV) Quantile regressions. To explore the heterogeneous effects of *EI* across the distribution of *GLT*, we perform quantile regressions at the 20th, 35th, 50th, 65th, and 80th percentiles. Columns (4–8) of Tab. 6 consistently demonstrate positive and statistically significant coefficients for *EI*, ranging from 0.155 (20th percentile, $p < 0.1$) to 0.266 (50th percentile, $p < 0.01$). These findings indicate that the positive relationship between *EI* and *GLT* is robust across varying levels of *GLT*, capturing both central and tail-end dynamics.

Overall, the results from these endogeneity tests provide strong evidence that the observed positive relationship between *EI* and *GLT* is not driven by endogeneity concerns. The consistent significance and magnitude of the coefficients across different estimation methods reinforce the validity and robustness of our primary conclusions.

3.6 Moderation analysis

To investigate the moderating role of *GRP* in the relationship between *EI* and *GLT*, interaction terms between *EI* and *GRP* were included in the analysis. Tab. 7 presents the results across two model specifications.

In Column (1), with only firm and year fixed effects, *EI* demonstrates a positive and statistically significant effect on *GLT* (coefficient = 0.206, $p < 0.01$), reaffirming its fundamental role in promoting green low-carbon outcomes. The direct effect of *GRP* on *GLT* is not statistically significant, suggesting that regulatory pressure alone does not directly drive *GLT*. However, the interaction term (*EI # GRP*) is negative and significant (coefficient = −0.010, $p < 0.05$), indicating that higher levels of *GRP* reduce the positive impact of *EI* on *GLT*. Also, in Column (2), after adding control variables, the results remain consistent. The coefficient for *EI* slightly decreases to 0.198 ($p < 0.01$), and the interaction term (*EI # GRP*) remains negative and significant (coefficient = −0.009, $p < 0.05$). A plausible explanation for these results might be that excessively high *GRP* could impose additional compliance burdens on firms or restrict their flexibility in implementing environmental innovations, thereby limiting their potential contribution to *GLT*.

The negative interaction effect observed in this study resonates with prior research highlighting the conditional impact of environmental regulations on innovation outcomes. For instance, Stojčić et al. (2024) propose that well-designed regulations can enhance competitiveness by incentivizing innovation, but overly restrictive frameworks may instead create inefficiencies. Similarly, Aragòn-Correa et al. (2019) argue that the relationship between regulation and innovation is contingent on the design and enforcement of policies, as excessive stringency may deter firms from pursuing innovative solutions due to increased costs or resource constraints.

In the context of *GLT*, these dynamics suggest that while regulatory pressure is essential for promoting sustainable practices, its intensity must be carefully calibrated. Regulations that are overly stringent may inadvertently hinder firms' ability to capitalize on the benefits of *EI*, particularly in industries where innovation cycles are long or require substantial investments. This highlights the importance of regulatory frameworks that strike a balance (providing sufficient incentives for *EI* without imposing excessive burdens that could dilute its effectiveness). The findings emphasize the nuanced interplay between *EI* and policy, suggesting that the success of *GLT* depends not only on the intrinsic quality of innovations but also on the external policy environment shaping their application.

Tab. 7: The moderating effect of *GRP* on the relationship between *EI* and *GLT*

Variables	(1)	(2)
	GLT	*GLT*
EI	0.206***	0.198***
	(0.047)	(0.049)
GRP	0.001	0.001
	(0.001)	(0.001)
EI # GRP	−0.010**	−0.009**
	(0.005)	(0.005)
Controls	No	Yes
Year FE	Yes	Yes
City FE	Yes	Yes
N	2,770	2,770
Adj. R^2	0.683	0.688

Note: Statistical significance is indicated by *, **, and *** for the 10, 5, and 1% levels, respectively; robustness standard errors are provided in brackets.

Source: own

Conclusions

Using data from 2012 to 2021 on 277 prefecture-level cities in China, this study investigates the relationship between environmental innovation and the green low-carbon transition, focusing on the moderating role of green regulatory pressure. The findings confirm that environmental innovation has a positive and significant impact on the green low-carbon transition, supporting the notion that innovative environmental practices are critical drivers of sustainability. Environmental innovation facilitates the adoption of low-carbon technologies, enhances resource efficiency, and reduces emissions, thereby accelerating the green low-carbon transition.

However, the results reveal that green regulatory pressure acts as a negative moderator in this relationship, diminishing the positive impact of environmental innovation on the green low-carbon transition. Excessive regulatory pressure can impose additional compliance burdens and limit firms' operational flexibility, constraining their ability to fully leverage environmental innovation. While regulation is necessary to encourage sustainable practices, overly stringent green regulatory frameworks may inadvertently hinder the contributions of environmental innovation to the green low-carbon transition.

From a theoretical perspective, this study integrates resource dependence theory and institutional theory to provide a nuanced understanding of the dynamics between environmental innovation and the green low-carbon transition under green regulatory pressure. Resource dependence theory underscores how environmental innovation enables firms to manage resource constraints and align operations with sustainability goals, reducing reliance on carbon-intensive resources. Institutional theory highlights the dual role of green regulatory pressure as both a driver and a constraint, compelling firms to adopt eco-friendly practices while potentially hindering their innovative capacity when overly restrictive. By demonstrating the contingent effects of regulatory intensity on innovation outcomes, this study advances both theories, emphasizing the need for adaptive regulatory frameworks that balance compliance imperatives with flexibility to foster sustainable transitions.

From a practical standpoint, these findings offer significant insights for policymakers and business leaders. Policymakers should design adaptive regulatory frameworks that encourage environmental innovation while avoiding excessive burdens on firms. This could involve, for example, implementing differentiated regulatory targets based on industry type and firm capacity, offering performance-based incentives for innovation outcomes, or strengthening coordination between environmental and industrial policy agencies to reduce regulatory fragmentation. Policies that provide flexibility, targeted incentives, and support for innovation can maximize the benefits of environmental innovation for achieving green low-carbon transitions. For firms, aligning innovation strategies with regulatory expectations while advocating for balanced policies is essential for sustaining progress toward low-carbon objectives.

In conclusion, this study provides empirical evidence of the positive role of environmental innovation in promoting the green low-carbon transition, while emphasizing the necessity of calibrated regulatory strategies. While environmental innovation serves as a key driver of sustainability, its effectiveness can be moderated by the intensity and design of green regulatory pressure. To build on these findings, future research may proceed along three promising directions. First, scholars could examine the long-term impacts of green regulatory pressure on different types of environmental innovation (such as incremental versus radical forms) and assess how these dynamics vary across sectors or national contexts. Such inquiry would deepen understanding of the institutional and economic conditions under which innovation most effectively contributes to sustainable transformation. Second, future studies may incorporate city-level contextual factors that influence firms' innovation capacity, including the presence of universities and research institutions, the educational attainment of the working-age population, the availability of sector-specific government technological support, and differences in industrial structure and energy intensity. Accounting for these heterogeneities may offer more nuanced insights into the drivers of environmental innovation and help explain spatial variations in green transition performance. Third, the measurement of green regulatory pressure could be enhanced by validating text-based indicators against concrete regulatory data (such as enforcement records, environmental expenditures, or pollution control efforts) in order to better align policy discourse with actual implementation.

References

Aftab, J., Abid, N., Cucari, N., & Savastano, M. (2022). Green human resource management and environmental performance: The role of green innovation and environmental strategy in a developing country. *Business Strategy and the Environment*, *32*(4), 1782–1798. https://doi.org/10.1002/bse.3219

Albitar, K., Borgi, H., Khan, M., & Zahra, A. (2022). Business environmental innovation and CO_2 emissions: The moderating role of environmental governance. *Business Strategy and the Environment*, *32*(4), 1996–2007. https://doi.org/10.1002/bse.3232

Ambec, S., Cohen, M. A., Elgie, S., & Lanoie, P. (2013). The Porter hypothesis at 20: Can environmental regulation enhance innovation and competitiveness? *Review of Environmental Economics and Policy*, *7*(1), 2–22. https://doi.org/10.1093/reep/res016

Anu, N., Singh, A. K., Raza, S. A., Nakonieczny, J., & Shahzad, U. (2022). Role of financial inclusion, green innovation, and energy efficiency for environmental performance? Evidence from developed and emerging economies in the lens of sustainable development. *Structural Change and Economic Dynamics*, *64*, 213–224. https://doi.org/10.1016/j.strueco.2022.12.008

Aragòn-Correa, J. A., Marcus, A. A., & Vogel, D. (2019). The effects of mandatory and voluntary regulatory pressures on firms' environmental strategies: A review and recommendations for future research. *Academy of Management Annals*, *14*(1), 339–365. https://doi.org/10.5465/annals.2018.0014

Baran, J., Szpor, A., & Witajewski-Baltvilks, J. (2020). Low-carbon transition in a coal-producing country: A labour market perspective. *Energy Policy*, *147*, 111878. https://doi.org/10.1016/j.enpol.2020.111878

Burdon, W. M., & Sorour, M. K. (2018). Institutional theory and evolution of 'a legitimate' compliance culture: The case of the UK financial service sector. *Journal of Business Ethics*, *162*(1), 47–80. https://doi.org/10.1007/s10551-018-3981-4

Cannavacciuolo, L., Ferraro, G., Ponsiglione, C., Primario, S., & Quinto, I. (2023). Technological innovation-enabling Industry 4.0 paradigm: A systematic literature review. *Technovation*, *124*, 102733. https://doi.org/10.1016/j.technovation.2023.102733

Feng, N., & Ge, J. (2024). How does fiscal policy affect the green low-carbon transition from the perspective of the evolutionary game? *Energy Economics*, *134*, 107578. https://doi.org/10.1016/j.eneco.2024.107578

Forés, B. (2024). Assessing the economic impacts of eco-innovation strategies across contexts shaped by different institutional and structural pressures. *Corporate Social Responsibility and Environmental Management*, *32*(1), 357–382. https://doi.org/10.1002/csr.2958

Gargallo, P., Lample, L., Miguel, J. A., & Salvador, M. (2024). Sequential management of energy and low-carbon portfolios. *Research in International Business and Finance*, *69*, 102263. https://doi.org/10.1016/j.ribaf.2024.102263

Ghisellini, P., Cialani, C., & Ulgiati, S. (2015). A review on circular economy: The expected transition to a balanced interplay of environmental and economic systems. *Journal of Cleaner Production*, *114*, 11–32. https://doi.org/10.1016/j.jclepro.2015.09.007

Jie, H., Zaman, S., Zaman, Q. uz, Shah, A. H., & Lou, J. (2023). A pathway to a sustainable future: Investigating the contribution of technological innovations, clean energy, and women's empowerment in mitigating global environmental challenges. *Journal of Cleaner Production*, *421*, 138499. https://doi.org/10.1016/j.jclepro.2023.138499

Khan, A. N., Jabeen, F., Mehmood, K., Ali Soomro, M., & Bresciani, S. (2023). Paving the way for technological innovation through adoption of artificial intelligence in conservative industries. *Journal of Business Research*, *165*, 114019. https://doi.org/10.1016/j.jbusres.2023.114019

Li, C., Sampene, A. K., Khan, A., & Wiredu, J. (2024). Environmental orientations to innovation ecosystem: The role of green transition, perceived environmental uncertainty and green transformational leadership. *Operations Management Research*, *18*(1), 48–71. https://doi.org/10.1007/s12063-024-00527-8

Lile, S., Ansari, S. (Shaz), & Urmetzer, F. (2024). Rethinking disruptive innovation: Unravelling theoretical controversies and charting new research frontiers. *Innovation*, 1–23. https://doi.org/10.1080/14479338.2024.2313197

Luo, K., Liu, Y., Chen, P.-F., & Zeng, M. (2022). Assessing the impact of digital economy on green development efficiency in the Yangtze River Economic Belt. *Energy Economics*, *112*, 106127. https://doi.org/10.1016/j.eneco.2022.106127

Mora-Contreras, R., Ormazabal, M., Hernández-Salazar, G., Torres-Guevara, L. E., Mejia-Villa, A., Prieto-Sandoval, V., & Carrillo-Hermosilla, J. (2023). Do environmental and cleaner production practices lead to circular and sustainability performance? Evidence from Colombian manufacturing firms. *Sustainable Production and Consumption*, *40*, 77–88. https://doi.org/10.1016/j.spc.2023.06.004

Owen, R., Brennan, G., & Lyon, F. (2018). Enabling investment for the transition to a low carbon economy: Government policy to finance early stage green innovation. *Current Opinion in Environmental Sustainability*, *31*, 137–145. https://doi.org/10.1016/j.cosust.2018.03.004

Porter, M. E., & Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. *The Journal of Economic Perspectives*, *9*(4), 97–118. https://doi.org/10.1257/jep.9.4.97

Pu, T. (2025). Does digital transformation propel innovation strategies? An empirical investigation based on machine learning and the dynamic GMM approach. *E&M Economics and Management*, *28*(2), 81–96. https://doi.org/10.15240/tul/001/2025-5-012

Pu, T., & Zulkafli, A. H. (2024). How does digital transformation affect innovation quality? *E&M Economics and Management*, *27*(4), 16–32. https://doi.org/10.15240/tul/001/2024-5-021

Stojčić, N., Vujanović, N., & Baum, C. F. (2024). Breaking or making futures: How laws and regulations shape innovation in emerging innovation systems. *Review of Managerial Science*, *19*(6), 1743–1782. https://doi.org/10.1007/s11846-024-00806-5

Többelmann, D., & Wendler, T. (2019). The impact of environmental innovation on carbon dioxide emissions. *Journal of Cleaner Production*, *244*, 118787. https://doi.org/10.1016/j.jclepro.2019.118787

Wang, C., Engels, A., & Wang, Z. (2017). Overview of research on China's transition to low-carbon development: The role of cities, technologies, industries and the energy system. *Renewable and Sustainable Energy Reviews*, *81*, 1350–1364. https://doi.org/10.1016/j.rser.2017.05.099

Wang, M., Li, Y., & Wang, Z. (2022). A nonlinear relationship between corporate environmental performance and economic performance of green technology innovation: Moderating effect of government market-based regulations. *Business Strategy and the Environment*, *32*(6), 3119–3138. https://doi.org/10.1002/bse.3290

Wu, J., Wu, Y., Guo, X., & Cheong, T. S. (2016). Convergence of carbon dioxide emissions in Chinese cities: A continuous dynamic distribution approach. *Energy Policy*, *91*, 207–219. https://doi.org/10.1016/j.enpol.2015.12.028

Xu, B. (2024). Environmental regulations, technological innovation, and low carbon transformation: A case of the logistics industry in China. *Journal of Cleaner Production*, *439*, 140710. https://doi.org/10.1016/j.jclepro.2024.140710

Xu, L., Fan, M., Yang, L., & Shao, S. (2021). Heterogeneous green innovations and carbon emission performance: Evidence at China's city level. *Energy Economics*, *99*, 105269. https://doi.org/10.1016/j.eneco.2021.105269

Xu, Y., Chin, W., Liu, Y., & He, K. (2022). Do institutional pressures promote green innovation? The effects of cross-functional coopetition in green supply chain management. *International Journal of Physical Distribution & Logistics Management*, *53*(7/8), 743–761. https://doi.org/10.1108/ijpdlm-03-2022-0104

Yang, G., Nie, Y., Li, H., & Wang, H. (2023). Digital transformation and low-carbon technology innovation in manufacturing firms: The mediating role of dynamic capabilities. *International Journal of Production Economics*, *263*, 108969. https://doi.org/10.1016/j.ijpe.2023.108969

York, J. G., & Venkataraman, S. (2010). The entrepreneur-environment nexus: Uncertainty, innovation, and allocation. *Journal of Business Venturing*, *25*(5), 449–463. https://doi.org/10.1016/j.jbusvent.2009.07.007

Yu, H., Wang, J., Hou, J., Yu, B., & Pan, Y. (2023). The effect of economic growth pressure on green technology innovation: Do environmental regulation, government support, and financial development matter? *Journal of Environmental Management*, *330*, 117172. https://doi.org/10.1016/j.jenvman.2022.117172

Zhao, Z., Zhao, Y., Lv, X., Li, X., Zheng, L., Fan, S., & Zuo, S. (2024). Environmental regulation and green innovation: Does state ownership matter? *Energy Economics*, *136*, 107762. https://doi.org/10.1016/j.eneco.2024.107762

Zhou, P., Ang, B. W., & Wang, H. (2012). Energy and CO_2 emission performance in electricity generation: A non-radial directional distance function approach. *European Journal of Operational Research*, *221*(3), 625–635. https://doi.org/10.1016/j.ejor.2012.04.022

Sickness presenteeism in remote work: Insights from a systematic review

***Katarzyna Wozniak-Jasinska*[1]**

[1] *Poznań University of Economics and Business, Institute of Economics, Department of Macroeconomics and Development Research, Poland, ORCID: 0000-0002-9077-328X, katarzyna.wozniak-jasinska@ue.poznan.pl (corresponding author).*

Abstract: *The increase in remote work, driven by the COVID-19 pandemic and technological advancements, has introduced numerous challenges for both employees and employers, including those related to health-related behaviors such as presenteeism. The aim of this paper is to identify the main determinants and reasons for presenteeism in remote and hybrid work, as well as key methodological aspects in researching this phenomenon. The study also highlights research gaps in the current body of knowledge and discusses implications based on the analysis of selected studies. The preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines and content analysis were applied to 425 papers retrieved from the Scopus database. The findings reveal that presenteeism increased during the COVID-19 pandemic. The main determinants of presenteeism include job demands, career pressure, access to sick leave, telework experience, overtime hours, indirect work control, employee health, supervisor support, detachment from work, and organizational culture. Specifically, presenteeism is negatively related to supervisor support and the ability to psychologically detach from work, while formal rules regarding presence and absence in the workplace play a role in explaining presenteeism in remote work environments. Most employees describe the home work environment as more "comfortable" for working while ill or unwell, which helps explain the higher level of presenteeism in remote work compared to on-site work. The main contributing factors included the ability to take breaks when needed, support from household members, and access to a better physical workspace. Finally, decisions to work while ill are shaped by factors such as how often one experiences illness, their general health condition, symptom severity and type, and personal illness perceptions. The findings underscore the need for flexible workplace policies and leadership approaches that promote employee health and support sustainable work practices.*

Keywords: *Employees health, hybrid work, work from home, working while ill.*

JEL Classification: *J22, J28, O33.*

APA Style Citation: Wozniak-Jasinska, K. (2026). Sickness presenteeism in remote work: Insights from a systematic review. *E&M Economics and Management*, *29*(2), 16–31. https://doi.org/10.15240/tul/001/2026-2-002

Introduction

The subsequent rise in remote working practices, resulting from technological development and the COVID-19 pandemic, has brought numerous challenges for both employees and employers, as well as several important advantages of this form of work. As part of the new world of work, the development of hybrid work, which combines remote work with on-site work, is also noticeable, and it is expected to become the most popular option after the COVID-19 pandemic (Kane et al., 2021). Generally, the positive aspects of work from home are evident in increased job productivity (Barrero et al., 2021); greater work and life satisfaction and various other benefits, such as time and cost savings, flexibility, including flexible working hours, work comfort, friendly work environment and greater job autonomy (Aksoy et al., 2022; Bavik et al., 2020; Nemțeanu et al., 2022). In contrast, remote work can also lead to deteriorating employee health, including ineffective communication, work-home interference, loneliness (Wang et al., 2021) and even the constant need to be online due to increased workload (Phadnis et al., 2021).

An important aspect of employees' health that changes during work from home is also presenteeism. Presenteeism refers to the behavior of working while being ill, serving as an alternative to being absent from work due to illness (Ruhle et al., 2020). The recent studies suggest that the consequences of presenteeism can be both positive (a functional behavior) and negative (a dysfunctional behavior). On one hand, presenteeism can have positive effects by supporting workplace inclusion of employees with chronic conditions (Karanika-Murray & Biron, 2020). On the other hand, it can lead to a decline in overall health when individuals continue to work despite being ill (Bergström et al., 2009). What is relevant, preliminary findings indicate that presenteeism is also common in remote work settings (Gerich, 2022).

Furthermore, some analyses indicate that work from home is a viable option for employees to continue working despite being ill (Nakrošienė et al., 2019). What is more, recent studies suggest that poor health, whether subjectively assessed or measured through specific health impairments, diminishes subjective well-being. Simultaneously, there are indirect effects whereby pursuing happiness serves as a buffer against bodily stress responses. However, these implications have less use in debilitating diseases (Binder & Buenstorf, 2018). Based on these implications, it is worth studying presenteeism in the context of remote and hybrid work more deeply. This paper is a systematic literature review, with the objective of categorizing and analyzing prior research on the presenteeism in remote and hybrid work. This systematic review draws on findings from prior research, identifies key aspects of research methodology related to presenteeism in remote work, outlines the main determinants of presenteeism in remote and hybrid work, highlights gaps in the current state of knowledge, and discusses the implications based on the analysis of selected studies.

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines and a content analysis were applied to 425 papers retrieved from the Scopus database. It also offers insights into future research to enhance the functional aspects of presenteeism while minimizing its negative consequences. This systematic review contributes to the literature by offering a structured overview of recent studies on presenteeism in remote and hybrid work and by identifying future research directions in this field. The originality of this research lies in not only highlighting authors' contributions but also systematizing findings on the prevalence of presenteeism during remote and hybrid work and the factors influencing this behavior.

The remainder of the paper is organized as follows: the second section outlines the research methodology, including the methods and materials used. The third section presents the results of the systematic review, covering descriptive statistics, the conceptualization of presenteeism definitions, measures of presenteeism, and the prevalence and factors influencing this behavior in remote work and implications for further research. The final section discusses the findings and implications for further research, followed by the recommendations for practice, and limitations of the study.

1 Theoretical background

1.1 Presenteeism as a decision making process

Given ongoing social, economic, and technological developments – along with the increase in remote work accelerated by the COVID-19 pandemic, presenteeism is understood as "*working while in a state of ill health*" (Ruhle et al., 2020, p. 346). What is more, presenteeism should not be limited to instances of acute illness alone; rather, it encompasses the broader behavior of continuing to work while experiencing any state of ill health (Ruhle et al., 2020).

Presenteeism results from conscious decision-making processes when individuals are faced with illness or poor health (Ruhle et al., 2020). In on-site work, employees who are sick must choose between two options: continuing to work by going to the office or staying home to recover (Karanika-Murray & Biron, 2020). However, with telework or remote work, the decision-making process becomes more complex. Employees now have a third option:

working from home despite feeling unwell (Ruhle et al., 2020). In such contexts, employees can choose whether to report sick, work from home while ill, or even attend the office despite their condition (Ruhle & Schmoll, 2021). Presenteeism, therefore, involves a complex decision-making process about whether to work while sick or to rest.

Generally, studies suggest that the consequences of presenteeism can be both positive (functional) and negative (dysfunctional). On the positive side, presenteeism can support workplace inclusion for employees with chronic conditions (Karanika-Murray & Biron, 2020). On the negative side, it can lead to a decline in overall health when individuals continue to work despite illness (Bergström et al., 2009). This raises important questions about presenteeism and its consequences in remote work settings and also underscores the need for an in-depth literature review. Therefore, this systematic review might highlight potential interventions that organizations and employees can use to mitigate the negative effects of presenteeism while supporting its functional aspects in remote work environments.

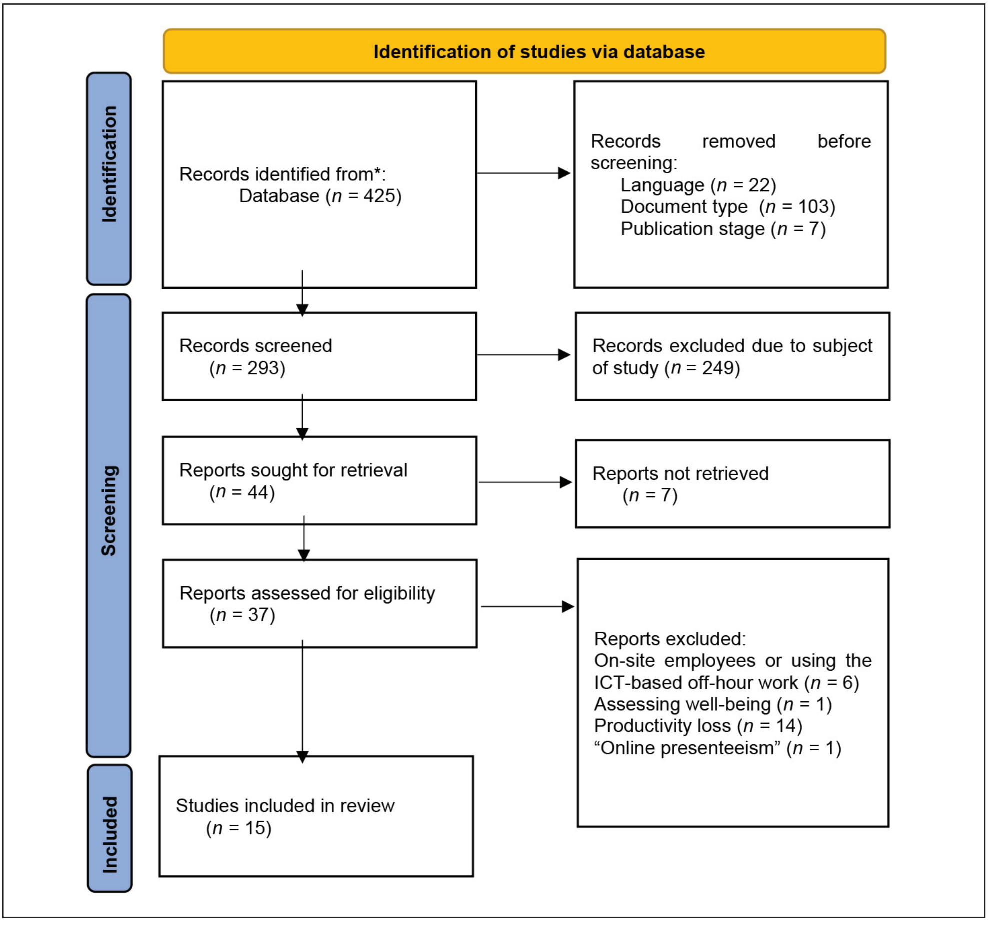

Fig. 1: PRISMA flow chart

Note: *n* – the number of papers.

Source: own

1.2 Conceptual foundations and formulation of research questions

The preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines were employed to systematically review, standardize, and report empirical results (Moher et al., 2009). PRISMA is a powerful tool for conducting systematic literature reviews because it enhances the quality of the review process from a methodological perspective (Arya et al., 2021). It also enables a structured and comprehensive approach. The PRISMA guidelines consist of four sequential processes: identification, screening, eligibility, and inclusion. All of these processes were strictly followed in this systematic review. The PRISMA procedure is illustrated in Fig. 1.

This systematic review addresses the following research questions:

RQ1: What are the definitions and measures of presenteeism during remote work?

RQ2: What are the main determinants of presenteeism in remote work?

RQ3: What are the major trends in studies on presenteeism and its consequences during remote work?

RQ4: What are the primary contributions of authors researching presenteeism in remote work, and what are the research gaps in this area?

By answering these questions, this review provides a comprehensive understanding of presenteeism in remote work, including its definitions, determinants, and consequences. It also offers insights for future research to enhance the functional aspects of presenteeism while minimizing its negative consequences. This systematic review contributes to the literature by offering a structured overview of recent studies on presenteeism in remote and hybrid work and by identifying future research directions in this field. The originality of this research lies in not only highlighting authors' contributions but also systematizing findings on the prevalence of presenteeism during remote and hybrid work and the factors influencing this behavior.

2 Research methodology

2.1 Data collection

A systematic review search was conducted in August 2024 using the electronic database Scopus. Elsevier's Scopus was chosen as the primary source of data for two main reasons. First, Scopus is the largest repository of citations and abstracts, including peer-reviewed research published in scientific journals, books, and conference proceedings (Elsevier, 2024). Second, Scopus has a broader citation index compared to the Web of Science, making it widely used for academic research (Zhao & Strotmann, 2015). While Google Scholar could yield more results, it may also include papers of local relevance, which could raise concerns about the quality of those articles (Piwońska et al., 2021). To obtain and explore the largest possible number of papers that assessed presenteeism in remote work, no specific time frame was set for the start of the search.

The following search terms were employed (TITLE-ABS-KEY("hybrid work*" OR "work* from home" OR "remote work*" OR "telework*" OR "telecommut*" OR "virtual work*" OR wfh) AND TITLE-ABS-KEY(presenteeism OR ill OR sick OR illness OR sickness OR unwell)). The particular keywords were selected to reflect the adopted research goals to the greatest extent possible. However, before developing the final combination of sets of keywords, the various combinations were carefully checked in order to choose the most efficient and suitable one, and, when necessary, the additional keywords were added. According to the previous analysis (e.g., Chain et al., 2019; Costa et al., 2019), documents were searched in article title, abstract, and author/indexed keywords of the papers. Firstly, the Scopus database returned 425 records. After preliminary screening of abstracts for further, more detailed analysis, 44 papers remained valid.

Selection criteria (inclusion and exclusion criteria)

In this study, the PCC (population (or participants)/concept/context) framework was employed in order to identify the main concepts in review questions and to ensure that any inclusion and exclusion criteria were not missed (Pollock et al., 2023). To be included in the systematic review, all papers were assessed based on the inclusion and exclusion criteria displayed in Tab. 1. The first criteria was the document type. According to this, papers that were not published in journals were excluded. Secondly, papers written in a language other than English were removed as well. Additionally, in order to ensure the transparency and credibility of the study, the analysis included articles that were finally published. To be eligible for inclusion in this systematic review, papers had to meet the following criteria: (1) participants were employees working part-time or full-time remotely; (2) the study analyzed presenteeism during work; (3) the study was empirical; and (4) it examined the impact of factors influencing presenteeism, as recommended by Johns (2010), Karanika-Murray and Cooper (2018).

Ruhle et al. (2020) studies that defined presenteeism as a loss of productivity were excluded from this literature review due to the fact that the loss of productivity due to illness resulting from working while unwell should not be strictly classified as presenteeism. Instead, they view it as a change in productivity resulting from presenteeism behavior (Johns, 2010; Karanika-Murray & Cooper, 2018; Ruhle et al., 2020). Additionally, one paper discussed the enforced constant presence and availability of employees during remote work, referred to by the authors as “online presenteeism”. However, since this presence was not associated with illness or feeling unwell, it does not meet the definition of presenteeism presented earlier and therefore cannot be considered presenteeism itself.

The review aimed to analyze empirical results from working contexts, specifically focusing on the working population. Therefore, studies involving student samples were excluded. However, if a study provided separate analyses of employees, it was included, but only the findings related to the working population were considered. Studies were excluded from the analysis if they: (1) included participants working only on-site; (2) referred to ICT-based off-hour work; (3) were review articles or similar; (4) discussed health issues unrelated to presenteeism; or (5) were written in a language other than English, were not published in journals, or were still in the publication process. If presenteeism was one of the factors examined (but not the primary focus of the analysis), the article was still included in the review as potentially important. However, the paper presents conclusions related to presenteeism specifically. A review of the full text resulted in 22 articles being excluded. Therefore, 15 papers were included in this review. All 15 papers included in this review were published in journals, the English language, between the years of 2002 and 2024.

Data extraction

Upon completing the review stages, including abstract screening and full-text reading, data extraction began for the included papers. The data were carefully extracted and the process was reviewed in consultation with two independent researchers in economics and management. Data extraction involved gathering information on various aspects relevant to the objective

Tab. 1: Outline of PCC framework for the review (inclusion and exclusion criteria)

PCC framework	Inclusion criteria	Exclusion criteria
Participants	Employees working part-time or full-time remotely	Unemployed people, students who do not work, employees working only on-site
Concept	Presenteeism during work	mWork (which refers to the ICT-based off-hour work) when employees normally work on-site
	Only empirical studies	Theoretical studies, review articles, editors introductions, etc.
	The impact of determinants shaping presenteeism	The impact of determinants shaping another phenomenon
Context	Remote or hybrid work environment	On-site work environment
	Papers written in English	Papers written in non-English
	Papers with final publication stage	Papers in press
	Texts published only in journals	Conference papers, book chapters, books, reviews, etc.

Source: own

of the systematic review, including the socio-demographics of respondents, data and methods used, study period, key findings on presenteeism and its influencing factors, as well as research gaps for future studies. Specifically, the following aspects were analyzed and presented into Microsoft Excel: (1) author(s); (2) year of publication; (3) title; (4) purpose; (5) study period; (6) model of work (hybrid, remote, on-site); (7) country/study location; (8) employment sectors/job type; (9) company size; (10) study population; (11) methods; (12) tools, data; (13) sample size; (14) COVID-19 context (pre-, during, post-pandemic); (15) definition of presenteeism; (16) measures of presenteeism; (17) the period of retrospective memory; (18) the prevalence of presenteeism; (19) list of factors shaping presenteeism; (20) main findings: factors shaping presenteeism; (21) research gaps.

Data coding and analysis

Data extraction was conducted to identify and assess key information on population demographics, methods, tools, and findings relevant to the implications of presenteeism in remote or hybrid work. Following the Braun and Clarke (2006) framework, a descriptive thematic analysis was performed. This process consists of six steps: the first step involved reviewing the data and formulating codes. Next, relevant data were collated for each code. In the following step, all relevant data were grouped into potential themes. After verifying the themes against the codes, all themes were refined. This procedure ensured the validity and robustness of the categories and themes.

3 Results and discussion

3.1 Statistical analysis of data resources

A total of 15 papers were included in the review. The analyzed articles were published in peer-reviewed journals, in English, between 2002 and 2024. Most articles were published in the years 2022 and 2023. These were publications written mainly in co-authorship, with international co-authorship accounting for almost 6.25%, and on average, with 19.5 citations per document.

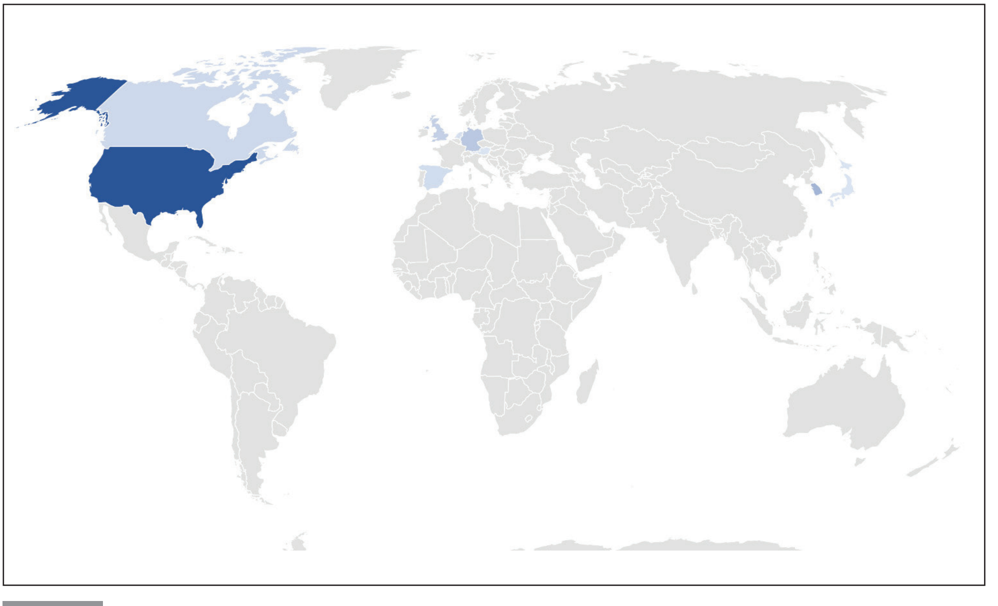

Fig. 2: Countries scientific production

Note: The data for this map were generated through “Biblioshiny”, a shiny app providing a web-interface for Bibliometrix software; different shades of blue indicate different productivity rate: dark blue – high productivity, grey – no paper.

Source: own (based on Scopus database)

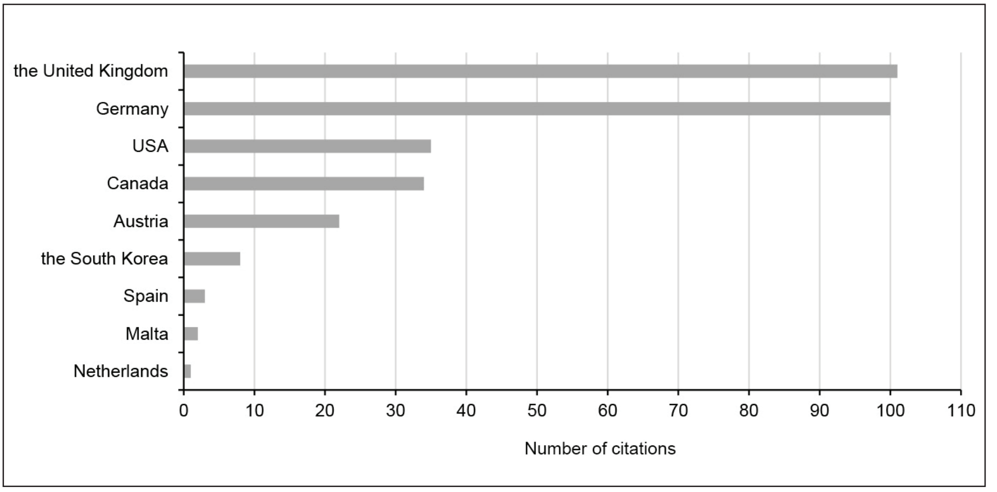

Fig. 3: Most cited countries

Source: own (based on Scopus database)

The geographical distribution of papers (based on all authors' affiliations) is concentrated in the South Korea, the Anglo-Saxon countries (USA, the UK, and Canada) and in other European countries (Germany, Spain, the Netherlands) (Fig. 2). The South Korea was the most productive country, probably due to the fact that cultural norms of dedication and work ethic (especially loyalty to one's workplace) play an important role in this country. Fig. 3 presents the 10 most cited countries according to the number of citations, with

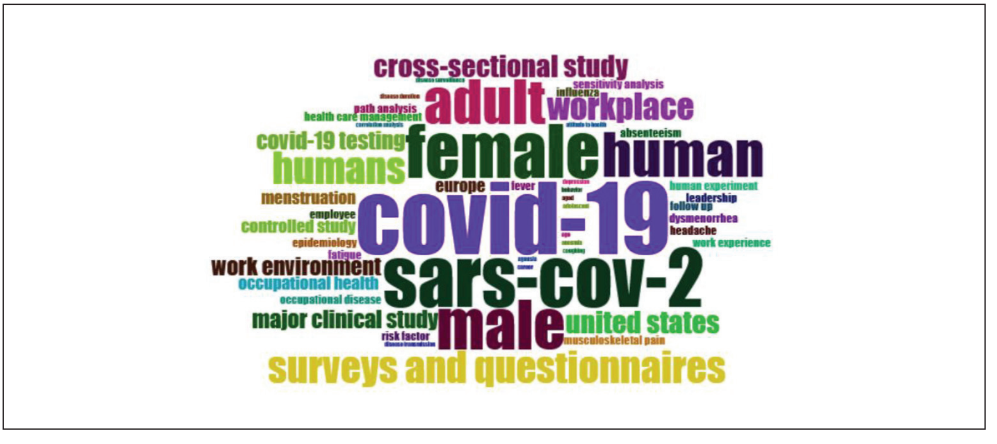

Fig. 4: Most frequent keywords used in the papers

Note: The word cloud was generated through "Biblioshiny", a shiny app providing a web-interface for Bibliometrix software; field: authors' keywords; the larger the size of the word, the more often it appeared in papers.

Source: own (based on Scopus database)

the United Kingdom having the most significant number of citations, followed by the Germany, the USA, and Canada.

The trending terms associated with presenteeism were analyzed to identify core thematic areas and to examine emerging patterns and developments in this field. The word cloud shows the most frequent keywords in the analyzed articles (Fig. 4). The larger the word, the more frequently the word appears in the analyzed articles. COVID-19 is the most emphasized in the literature reviewed, followed by sars-cov-2, female, male, adult, humans, surveys and questionnaires, workplace, cross-sectional study, work environment, occupational health, absenteeism, etc. Over recent years, the issue of presenteeism has gained increasing attention, exacerbated by the effects of COVID-19 pandemic.

3.2 Definitions and conceptual approaches to presenteeism in remote work

Studies on presenteeism while working from home use various definitions (Tab. 2). Generally, based on the definitions collected from research on presenteeism in remote or hybrid work settings, presenteeism is usually defined as the behavior of employees who work despite being sick or unwell. Most studies identify presenteeism as a phenomenon where employees continue to work (remotely, usually at home) even when they are ill (e.g., Fiorini, 2024; Gerich, 2022; Goñi-Legaz et al., 2024; Hadjisolomou et al., 2022; Magalhães et al., 2022; Probst et al., 2021; Ruhle & Schmoll, 2021; Ryoo et al., 2023). This understanding of presenteeism, as working while sick, aligns with definitions that emphasize it as a decision, whether conscious or complex decision-making, where the employee chooses to work despite illness symptoms (Hadjisolomou et al., 2022; Schmitz et al., 2023) or continues to engage in work activities despite the presence of illness symptoms (Cook & Van Den Hoek, 2023).

3.3 Measures of presenteeism in remote work

A systematic review highlights significant variations in the measurement of presenteeism (Tab. 3). Most studies focus on the extent and likelihood of presenteeism resulting from being sick, feeling unwell, or experiencing pain. Specifically, some measures focus solely on the extent of presenteeism, studying its prevalence (Ryoo et al., 2023, 2024) or duration, with most considering the number of days (e.g., Ahmed et al., 2020; Schmitz et al., 2023; Shafer et al., 2023), although some also take into account the number of occurrences (Fiorini, 2024). Several measures include different items to assess presenteeism. For instance, Cook and Van Den Hoek (2023) consider six items related not only to working despite having symptoms or pain but also to working even though a doctor

Tab. 2: The definitions of presenteeism in remote work

Definition	Source
A conscious decision to work when ill (even if employees only work for specific tasks/meetings or a shorter period than they contractually agreed)	(Schmitz et al., 2023)
Workers' decisions to work when sick	(Hadjisolomou et al., 2022)
The behaviour of continuing to engage in work activities despite the presence of symptoms" and the result of a complex decision-making process	(Cook & Van Den Hoek, 2023)
Going to work/working/continue working despite being sick/ill	(Fiorini, 2024; Gerich, 2022; Goñi-Legaz et al., 2024; Magalhães et al., 2022; Probst et al., 2021; Ruhle & Schmoll, 2021; Ryoo et al., 2023, 2024; Spinks, 2002; Steidelmüller et al., 2020)
The state of attending work when individual is unwell or the act of working in a state of ill-health	(Biron et al., 2021)

Source: own

advised against it, as well as taking medication in order to continue working. Furthermore, Steidelmüller et al. (2020) and Gerich (2022), in addition to assessing the prevalence and duration of presenteeism (in days), also employed a measure of sickness presenteeism propensity. Probst et al. (2021), focusing on the COVID-19 perspective, considered the duration of presenteeism (occurrences) in three different areas: feeling under the weather, possible exposure to someone with COVID-19, and being diagnosed with COVID-19. Meanwhile, Ruhle and Schmoll (2021), using open-ended questions, concentrated on the reasons and main characteristics of presenteeism, such as organizations' attitudes toward employee attendance and absences in the event of illness, or differences in how working while sick was treated before the crisis. Additionally, Fiorini (2024) used a single-item measure to assess the instances when an employee worked while feeling unwell, incorporating the role of sick leave into this measure, through the question *"How many times had you worked, from their workplace or remotely, despite feeling sufficiently unwell to take sick leave during the first 12 months of the pandemic?"* (Tab. 3).

Moreover, the different periods of retrospective memory reported by employees were considered. Most studies referred to a retrospective memory period of 12 months (Fiorini, 2024; Gerich, 2022; Goñi-Legaz et al., 2024; Probst et al., 2021; Ryoo et al., 2023; 2024; Steidelmüller et al., 2020). Other perspectives included a retrospective memory period of 6 months (Cook & Van Den Hoek, 2023), 3 months (Ruhle & Schmoll, 2021; Schmitz et al., 2023), 28 days (Shafer et al., 2023), 7 days (Biron et al., 2021), or even the first days of illness (Ahmed et al., 2020). Generally, authors who referred to periods shorter than 12 months indicated that they were trying to reduce recall bias (e.g., Ruhle & Schmoll, 2021). Nevertheless, a 12-month recall period is often used in studies on presenteeism and working from home (telework).

Data in most studies were collected through the authors' own surveys and sometimes interviews (e.g., Biron et al., 2021; Fiorini, 2024; Gerich, 2022; Schmitz et al., 2023). Some papers also used data collected from national or international surveys conducted by international institutions or specific countries, e.g., the European Working Conditions Survey (Goñi-Legaz et al., 2024; Steidelmüller et al., 2020) or the 6th Korean Working Conditions Survey (Ryoo et al., 2023; 2024). In closing, most presenteeism measures rely on subjective

Tab. 3: The measures of presenteeism in remote work – Part 1

Measure	The period of retrospective memory	The data collected	Source
The question inquiring about working while ill (the response "yes" was considered to be the presenteeism)	12 months ago	The 6th Korean Working Conditions Survey	(Ryoo et al., 2023; 2024)
The question *"How many times had you worked, from their workplace or remotely, despite feeling sufficiently unwell to take sick leave during the first 12 months of the pandemic*?"	12 months ago	Authors' survey	(Fiorini, 2024)
The question "*On how many days did you work remotely in the last 3 months although you felt ill*?"	3 months ago	Authors' survey	(Schmitz et al., 2023)
The number of days worked during illness by summing the days worked onsite and the days teleworked during illness	28 days ago	Authors' survey	(Shafer et al., 2023)
The question "*In the last week (7 days), how many days did you work while you had a health problem?*"; where health problems are any physical or emotional problem or symptom	7 days ago	Authors' survey	(Biron et al., 2021)

Tab. 3: The measures of presenteeism in remote work – Part 2

Measure	The period of retrospective memory	The data collected	Source
The question for sickness presence days: "*Approximately how many days did you work during the past 12 months even when your health state would have justified taking sick leave?*"; presenteeism was calculated as the number of sickness presence days divided by the number of health events, whereby the number of health events is the sum of sickness presence and absence days (for both sickness presence and sickness absence, responses of more than 60 days were removed to avoid bias due to long-term sickness)	12 months ago	Authors' survey	(Gerich, 2022)
The six items: (1) "*I appeared at work despite having symptoms/pain*"; (2) "*I worked even though my doctor advised me not to*"; (3) "*I worked even though I experienced severe illness symptoms/pain*"; (4) "*I worked the entire day/shift even though I was experiencing symptoms/pain*"; (5) "*I took medication to manage acute symptoms/pain in order to be able to work*"; (6) "*I dragged myself to work even though I was experiencing symptoms/pain*"	6 months ago	Authors' survey	(Cook & Van Den Hoek, 2023)
The three questions: (1) "*Since the start of the pandemic, how often have you gone to work despite feeling under the weather (e.g., cough, low fever, sore throat, fatigued, etc.)?*"; (2) "*Since the start of the pandemic, how often have you gone to work despite being possibly exposed to someone with COVID-19*"; (3) *"Since the start of the pandemic, how often have you gone to work despite being diagnosed with COVID-19*"	12 months ago	Authors' survey	(Probst et al., 2021)
Work attendance during the first 3 days of illness (including the number of days worked at the usual workplace and number of days worked from home)	The first days of illness	Authors' survey	(Ahmed et al., 2020)
Three different approaches: (1) the question "*Over the past 12 months did you work when you were sick?*" (variable binary); (2) sickness presenteeism days, i.e., the number of working days showing presenteeism (individuals reported to have not worked when sick or not having sick were coded with a zero and were excluded); (3) the sickness presenteeism propensity, calculated as the ratio of days showing presenteeism and the sum of days in presenteeism and sickness absence, and it ranges between 0 (i.e., all days of sickness were spent in sickness absence) and 1 (i.e., all days of sickness were days showing presenteeism)	12 months ago	The European Working Conditions Survey (EWCS) 2015	(Steidelmüller et al., 2020; Goñi-Legaz et al., 2024)

Tab. 3: The measures of presenteeism in remote work – Part 3

Measure	The period of retrospective memory	The data collected	Source
The opened questions: (Q1) "*Please describe why you decided to work while sick. What were the reasons for this? How did this behavior differ from your behavior before the pandemic?*"; (Q2) "*Where were you working at home, the office, or somewhere else? What were the hours of work? Were there any special circumstances?*"; (Q3) "*Please describe the extent to which COVID-19 has led your organization to address employee attendance and absences in the event of illness. Briefly state whether you have been treated differently in the context of working with an illness then you were before the crisis? If so, how?*"	3 months ago	Authors' survey	(Ruhle & Schmoll, 2021)
The percentage of employees report work despite having health conditions that affected their job performance, and chose not to seek treatment	NA	Authors' survey	(Spinks, 2002)

Source: own

data, such as the number of times or days employees feel sick, ill, or unwell but still decide to continue working. As such, they are largely based on employees' personal perceptions and self-reports. These measures typically do not indicate whether the illness or poor health was confirmed by a doctor. However, some studies do incorporate additional aspects such as sick leave or medical consultations. For example, the study by Cook and Van Den Hoek (2023) examines whether an individual worked despite receiving medical advice not to.

3.4 A thematic discussion of presenteeism influences and emerging research needs

Based on the literature review, two main groups of factors influencing presenteeism in remote work can be distinguished: socio-demographic and job-related. The most frequently indicated socio-demographic factors include age, gender, education, marital status, having children under 18 years of age, caregiving responsibilities for children, grandchildren, disabled relatives, or the elderly (at least several times a week), income, household size, and race/ethnicity. The most commonly cited job-related characteristics are enterprise size, job type/occupation, employment sector, type of work contract, leadership role, work experience, telework intensity, psychological demands, overtime hours, job control, psychosocial safety climate, telework experience, detachment from work, supervisor support, organizational policies, employee attitudes toward CDC COVID-19 prevention guidelines, and work-family conflict. Furthermore, some studies also explored specific reasons for presenteeism in remote work.

Although the studies were diverse, some overlapping aspects were identified in the relationship between presenteeism and working from home, as well as the factors shaping this behavior during telework. Most studies indicate that presenteeism is positively associated with home-based telework (e.g., Goñi-Legaz et al., 2024; Ryoo et al., 2023, 2024) and that this behavior increased during the COVID-19 pandemic (e.g., Ruhle & Schmoll, 2021). Regarding moderators, most studies suggest that presenteeism is more frequent among women and older employees (e.g., Fiorini, 2024; Goñi-Legaz et al., 2024). However, some studies found no statistically significant differences in presenteeism based on employee characteristics (e.g., Ahmed et al., 2020; Biron et al., 2021; Ryoo et al., 2023).

Moreover, previous studies indicate that presenteeism is positively associated with several factors, including: job demands (Biron et al., 2021; Ryoo et al., 2024), career pressure (Ruhle & Schmoll, 2021), working in larger companies (Schmitz et al., 2023), holding a managerial role (Ruhle & Schmoll, 2021), the lack of access to paid sick leave (Ahmed et al., 2020), prior telework experience (Shafer et al., 2023), utilizing telework for work intensification, overtime working hours, indirect work control (Gerich, 2022), employees' poor health (Cook & Van Den Hoek, 2023; Goñi-Legaz et al., 2024). Conversely, presenteeism is negatively associated with supervisor support (Schmitz et al., 2023), positive attitudes toward COVID-19 prevention at workplace (Probst et al., 2021), or detachment from work (Schmitz et al., 2023). Several studies have also highlighted various reasons contributing to presenteeism in remote work, pointing to factors such as work demands, health conditions, illness perceptions (e.g., Fiorini, 2024; Ryoo et al., 2024), and formal policies on attendance and absence (Ruhle & Schmoll, 2021). The detailed results of the analyzed papers are available in the Supplementary materials.

Furthermore, the main reasons for presenteeism primarily fall into three categories: work and job-related, the characteristics of remote work environment, and health-related factors. Work and job-related factors include the casual nature of his contract, and related with it job insecurity and financial instability (Hadjisolomou et al., 2022) the difficult to finding a replacement at workplace, career pressure (Ruhle, & Schmoll, 2021), heavy workloads (Fiorini, 2024), professional obligation to co-workers (Ahmed et al., 2020), and official rules and regulations about absence and presence and organizations (Hadjisolomou et al., 2022; Ruhle & Schmoll, 2021). Key characteristics of the remote work environment (usually when working from home) are the lack of commuting (Fiorini, 2024; Ruhle & Schmoll, 2021; Spinks, 2002), and the general possibility of adjustment of work volume during work from home or to adjust work for specific illnesses (Fiorini, 2024; Ruhle & Schmoll, 2021). Specifically, employees usually assessed the home work environment as more "comfortable" for working when ill or unwell. This was due to the ability to take breaks when needed, receive support from others in the household, and benefit from a better physical workspace (e.g., proper lighting, ergonomic desk/chair setup) (Fiorini, 2024; Spinks, 2002). Health- and illness-related reasons for presenteeism include disease frequency, general health status, illness perceptions, and symptom types (Fiorini, 2024). For instance, some employees reported presenteeism because their symptoms were mild enough not to require sick leave, and they believed these symptoms had little or no impact on their work performance (Ahmed et al., 2020; Fiorini, 2024). Additionally, presenteeism was reported in relation to specific conditions such as pain, migraines, and mental health symptoms. Some employees also admitted working while sick due to feelings of guilt about using sick leave (Fiorini, 2024).

Regarding research gaps, several studies note the need for research on possible moderators and mediators of presenteeism, especially job-related characteristics, such as company characteristics (e.g., enterprise size), workplace policies and culture, colleague support, occupation, and availability of paid sick leave (Ahmed et al., 2020; Biron et al., 2021; Schmitz et al., 2023; Shafer et al., 2023), as well as factors related to health status, e.g., physical exercise habits (Ryoo et al., 2024). Another identified research gap highlights the lack of investigation into the impact of blurred boundaries while working remotely on sickness presenteeism (Biron et al., 2021). Additionally, there is still a lack of research on the reasons for presenteeism in remote work settings (Schmitz et al., 2023).

Authors also highlight that the role of the consequences of sickness presenteeism (both positive and negative) has not been sufficiently explored. Specifically, there is still insufficient research on the effects of presenteeism (especially positive effects) during telework (Ruhle & Schmoll, 2021), as well as its effects on productivity due to working while ill (particularly for individuals with minor illnesses) and on measuring output while working when ill or unwell (Ahmed et al., 2020; Fiorini, 2024; Shafer et al., 2023). Several studies also indicate the need for research on the circumstances of health management in virtual work environments (Spinks, 2002), and on the relationship between presenteeism and the diversity, duration, and occurrence of specific symptoms, for example, the menstrual cycle (Cook & Van den Hoek, 2023).

Regarding methodological issues, previous studies note the need for longitudinal research, applying prospective research designs or diary studies to explore sickness presenteeism in remote and hybrid work (Ahmed et al., 2020; Gerich, 2022) or cohort studies (Ryoo et al., 2023). Finally, authors highlight the need to examine and compare sickness presenteeism across different work settings, i.e., on-site and remote work (Goñi-Legaz et al., 2023; Ryoo et al., 2023). Therefore, recommendations for future research include examining the impact of job-related characteristics on attending work while ill, especially in terms of workplace policies and organizational culture. A growing body of scientific evidence on the impact of workplace policies and culture on presenteeism in remote and hybrid work should also encourage more in-depth research into contextual aspects of presenteeism, such as variations in presenteeism climates, norms, and organizational cultures across occupations and sectors. Additionally, considering that one of the defining characteristics of remote work is the weakening, or even absence, of physical and temporal boundaries between work and non-work domains, future studies should explore how these blurred boundaries influence employee health and performance. It is also essential for future research to assess employee productivity while working during illness across different organizational settings, i.e., in remote and on-site work. Regarding methodological issues in future studies on sickness presenteeism in remote work, research should pay special attention to conducting longitudinal studies, including diary studies. Researchers should also focus on cohort and intervention studies on presenteeism and its related occupational health risks in the context of remote work. Future studies should not rely solely on retrospective recall to study sickness presenteeism but should also employ hypothetical scenarios to examine sickness presenteeism propensity, especially from a long-term perspective. Finally, while most studies focus only on remote work, evidence on hybrid work arrangements remains an interesting area for further investigation.

Conclusions

This systematic review reveals several key findings. First, most studies define presenteeism as the behavior of working while ill or unwell, with definitions slightly varying across disciplinary contexts. Most empirical studies were conducted in the South Korea, the USA, the United Kingdom, Canada, and the European countries, primarily focusing on the prevalence of presenteeism (commonly understood as working or continuing to work despite being sick or unwell) and the factors influencing this behavior. Additionally, it is worth noting that studies on remote work often portray presenteeism as a negative phenomenon, frequently overlooking its potential functional aspects.

Furthermore, the diversity of eligible studies was also reflected in the types of measurements, recall periods, and data sources used. Regarding the prevalence of presenteeism, most studies measured it as the number of days or instances when employees felt sick, ill, or unwell but chose to continue working. Notably, some studies also explored related aspects, such as sick leave or doctor visits. For example, the study by Cook and Van Den Hoek (2023) examined whether individuals worked despite their doctor's advice not to. Most studies measuring the prevalence of presenteeism referred to a retrospective memory period of 12 months (e.g., Fiorini, 2024; Gerich, 2022; Probst et al., 2021; Ryoo et al., 2023, 2024; Steidelmüller et al., 2020). Other studies used shorter periods, such as 6 months (Cook & Van Den Hoek, 2023), 3 months (e.g., Ruhle & Schmoll, 2021; Schmitz et al., 2023), 28 days (Shafer et al., 2023), 7 days (Biron et al., 2021), or even just the first days of illness (Ahmed et al., 2020). Authors using shorter recall periods often aimed to reduce recall bias (e.g., Ruhle & Schmoll, 2021). Nonetheless, a 12-month recall period remains rather common in studies on presenteeism and working from home. Data in most studies were collected through the authors' surveys or, occasionally, interviews. Some papers also utilized data from national or international surveys conducted by institutions or specific countries (e.g., the European Working Conditions Survey).

The findings show that presenteeism increased during the COVID-19 pandemic. Key drivers of presenteeism include job demands, career pressure, lack of access to paid sick leave, prior telework experience, overtime hours, and indirect work control. Furthermore, employees working in larger companies or holding managerial positions also reported higher levels of presenteeism. Previous research also reveals that presenteeism is negatively related to supervisor support and the ability to psychologically detach from work. Additionally, formal rules regarding presence and absence in the workplace play a role in explaining presenteeism in remote work environments. Moreover, most employees described the home work environment as more "comfortable" for working while ill or unwell. The main contributing factors included the ability to take breaks when needed, support from household members, and access to a better physical workspace. Finally, decisions to work while ill are shaped by factors such as how often one experiences illness, their general health condition, symptom severity and type, and personal illness perceptions.

The findings from this systematic review have significant implications for practitioners, policymakers, and organizations. Specifically, they highlight the determinants and consequences of presenteeism in remote and hybrid work, which can contribute to more effective management strategies in these settings. First, it should be emphasized that presenteeism in remote and hybrid work can lead to both positive and negative outcomes. From the employee's perspective, presenteeism may offer an opportunity to continue working during illness or reduced well-being (e.g., when symptoms are relatively mild), particularly because work can be performed in a more comfortable, home-like environment. However, it must also be considered that presenteeism, especially when driven by external pressures such as high job demands or the need for overtime, can result in negative outcomes, including increased stress and deterioration of health, particularly over the long term. From the employer's perspective, presenteeism may help maintain productivity if employees continue working despite illness or poor well-being. It may also alleviate operational challenges, such as finding replacements for absent staff. However, it can also reflect issues such as low managerial support, which may, over time, negatively impact organizational performance. Therefore, organizations should implement flexible work policies that enhance employee autonomy while minimizing the negative consequences of working while ill. Based on the available evidence and global trends in the development of remote work practices, companies should be encouraged to design health-supportive policies that empower employees to make sustainable decisions about working during illness. Additionally, it would be beneficial to develop specific guidelines for remote and hybrid work that support both employees and their leaders. Providing training for managers in trust-based leadership styles may also improve the management of employee health in remote workplaces. These interventions can not only support employees in making sustainable decisions regarding presenteeism but also foster a healthier and more resilient remote work culture.

Acknowledgments: *The research was financed by the Poznań University of Economics and Business as part of a subsidy received from the Ministry of Science and Higher Education. The author gratefully acknowledges Professor Sandra Ohly and her research team, as well as Dr. Łukasz Kryszak for their invaluable guidance and support in the process of preparing this paper.*

References

Ahmed, F., Kim, S., Nowalk, M. P., King, J. P., VanWormer, J. J., Gaglani, M., Zimmerman, R. K., Bear, T., Jackson, M. L., Jackson, L. A., Martin, E., Cheng, C., Flannery, B., Chung, J. R., & Uzicanin, A. (2020). Paid leave and access to telework as work attendance determinants during acute respiratory illness, United States. *Emerging Infectious Diseases*, *26*(1), 26–33. https://doi.org/10.3201/eid2601.190743

Aksoy, C. G., Barrero, J. M., Bloom, N., Davis, S. J., Dolls, M., & Zarate, P. (2023). *Working from home around the globe* [report No. 53]. EconPol Policy Brief. https://www.ifo.de/en/econpol/publications/2023/working-paper/working-home-around-globe-2023-report

Arya, S., Kaji, A. H., & Boermeester, M. A. (2021). PRISMA reporting guidelines for meta-analyses and systematic reviews. *JAMA Surgery*, *156*(8), 789–790. https://doi.org/10.1001/jamasurg.2021.0546

Barrero, J. M., Bloom, N., & Davis, S. (2021). *Why working from home will stick* [working paper No. 28731]. National Bureau of Economic Research. https://doi.org/10.3386/w28731

Bavik, Y. L., Shaw, J. D., & Wang, X.-H. (Frank). (2020). Social support: Multidisciplinary review, synthesis, and future agenda. *Academy* *of Management Annals*, *14*(2), 726–758. https://doi.org/10.5465/annals.2016.0148

Bergström, G., Bodin, L., Hagberg, J., Aronsson, G., & Josephson, M. (2009). Sickness presenteeism today, sickness absenteeism tomorrow? A prospective study on sickness presenteeism and future sickness absenteeism. *Journal of Occupational and Environmental Medicine*, *51*(6), 629–638. https://doi.org/10.1097/JOM.0b013e3181a8281b

Binder, M., & Buenstorf, G. (2018). Smile or die: Can subjective well-being increase survival in the face of substantive health impairments? *Economics & Human Biology*, *31*, 209–227. https://doi.org/10.1016/j.ehb.2018.08.004

Biron, C., Karanika-Murray, M., Ivers, H., Salvoni, S., & Fernet, C. (2021). Teleworking while sick: A three-wave study of psychosocial safety climate, psychological demands, and presenteeism. *Frontiers in Psychology*, *12*, 1–14. https://doi.org/10.3389/fpsyg.2021.734245

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Chain, C. P., Santos, A. C., Castro, L. G., & Prado, J. W. (2019). Bibliometric analysis of the quantitative methods applied to the measurement of industrial clusters. *Journal of Economic Surveys*, *33*(1), 60–84. https://doi.org/10.1111/joes.12267

Cook, A., & Van Den Hoek, R. (2023). Period pain presenteeism: Investigating associations of working while experiencing dysmenorrhea. *Journal of Psychosomatic Obstetrics & Gynecology*, *44*(1), 1–12. https://doi.org/10.1080/0167482X.2023.2236294

Costa, D. F., de Melo Carvalho, F., & de Melo Moreira, B. C. (2019). Behavioral economics and behavioral finance: A bibliometric analysis of the scientific fields. *Journal of Economic Surveys*, *33*(1), 3–24. https://doi.org/10.1111/joes.12262

Elsevier. (2024). *Scopus: Comprehensive, multidisciplinary, trusted abstract and citation database*. https://www.elsevier.com/products/scopus

European Commission. (2022). *Digital Economy and Society Index (DESI)* [report]. https://digital-strategy.ec.europa.eu/en/library/digital-economy-and-society-index-desi-2022

Fiorini, L. A. (2024). Remote workers' reasons for changed levels of absenteeism, presenteeism and working outside agreed hours during the COVID-19 pandemic. *Sage Open*, *14*(1), 1–18. https://doi.org/10.1177/21582440241240636

Gerich, J. (2022). Home-based telework and presenteeism: New lessons learned from the Covid-19 pandemic. *Journal of Occupational and Environmental Medicine*, *64*(3), 243–249. https://doi.org/10.1097/JOM.0000000000002414

Goñi-Legaz, S., Núñez, I., & Ollo-López, A. (2024). Home-based telework and job stress: The mediation effect of work extension. *Personnel Review*, *53*(2), 545–561. https://doi.org/10.1108/PR-02-2022-0111

Hadjisolomou, A., Mitsakis, F., & Gary, S. (2022). Too scared to go sick: Precarious academic work and 'presenteeism culture' in the UK higher education sector during the Covid-19 pandemic. *Work, Employment and Society*, *36*(3), 569–579. https://doi.org/10.1177/09500170211050501

Johns, G. (2010). Presenteeism in the workplace: A review and research agenda. *Journal of Organizational Behavior*, *31*(4), 519–542. https://doi.org/10.1002/job.630

Kane, G. C., Nanda, R., Phillips, A., & Copulsky, J. (2021). Redesigning the post-pandemic workplace. *MIT Sloan Management Review*, *62*(3), 12–14.

Karanika-Murray, M., & Biron, C. (2020). The health-performance framework of presenteeism: Towards understanding an adaptive behaviour. *Human Relations*, *73*(2), 242–261. https://doi.org/10.1177/0018726719827081

Karanika-Murray, M., & Cooper, C. L. (2018). Presenteeism: An introduction to a prevailing global phenomenon. In L. Lu & C. L. Cooper (Eds), *Presenteeism at work* (pp. 9–34). Cambridge University Press. https://doi.org/10.1017/9781107183780.003

Magalhães, S., Barbosa, J., & Borges, E. (2022). Presenteeism in non-academic staff in a public university context: Prevalence, associated factors, and reasons to work while sick during the COVID-19 pandemic. *International Journal of Environmental Research and Public Health*, *19*(22), 14966. https://doi.org/10.3390/ijerph192214966

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *7*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Nakrošienė, A., Bučiūnienė, I., & Goštautaitė, B. (2019). Working from home: Characteristics and outcomes of telework. *International Journal of Manpower*, *40*(1), 87–101. https://doi.org/10.1108/IJM-07-2017-0172

Nemțeanu, M. S., Dinu, V., Pop, R. A., & Dabija, D. C. (2022). Predicting job satisfaction and work engagement behavior in the COVID-19 pandemic: A conservation of resources theory approach. *E&M Economics and Management*, *25*(2), 23–40. https://doi.org/10.15240/tul/001/2022-2-002

Phadnis, S., Sengupta, S., & Chakraborty, A. (2021). Work from home, mental health and employee needs: A pilot study in selected information technology organizations in India. *Asia Pacific Journal of Health Management*, *16*(3), 103–110. https://doi.org/10.24083/apjhm.v16i3.977

Piwońska, K., Urbańska, M., Pilarska, E., Miniszewski, M., Staniszewski, J., & Kryszak, Ł. (2021). Managing sustainable consumption of durable goods – A systematic literature review. *Management*, *25*(2), 73–90. https://doi.org/10.2478/manment-2019-0074

Pollock, D., Peters, M. D. J., Khalil, H., McInerney, P., Alexander, L., Tricco, A. C., Evans, C., de Moraes, É. B., Godfrey, C. M., Pieper, D., Saran, A., Stern, C., & Munn, Z. (2023). Recommendations for the extraction, analysis, and presentation of results in scoping reviews. *JBI Evidence Synthesis*, *21*(3), 520–532. https://doi.org/10.11124/JBIES-22-00123

Probst, T. M., Lee, H. J., Bazzoli, A., Jenkins, M. R., & Bettac, E. L. (2021). Work and non-work sickness presenteeism: The role of workplace COVID-19 climate. *Journal of Occupational and Environmental Medicine*, *63*(8), 713–718. https://doi.org/10.1097/JOM.0000000000002240

Ruhle, S. A., Breitsohl, H., Aboagye, E., Baba, V., Biron, C., Correia Leal, C., Dietz, C., Ferreira, A. I., Gerich, J., Johns, G., Karanika-Murray, M., Lohaus, D., Løkke, A., Lopes, S. L., Martinez, L. F., Miraglia, M., Muschalla, B., Poethke, U., Sarwat, N., Schade, H., Steidelmuller, C., Vinberg, S., Whysall, Z., & Yang, T. (2020). "To work, or not to work, that is the question" – Recent trends and avenues for research on presenteeism. *European Journal of Work and Organizational Psychology*, *29*(3), 344–363. https://doi.org/10.1080/1359432X.2019.1704734

Ruhle, S. A., & Schmoll, R. (2021). COVID-19, telecommuting, and (virtual) sickness presenteeism: Working from home while ill during a pandemic. *Frontiers in Psychology*, *12*, 1–13. https://doi.org/10.3389/fpsyg.2021.734106

Ryoo, S.-W., Min, J.-Y., Lee, D.-W., Choi, B.-Y., Choi, J., Kim, H.-Y., & Min, K.-B. (2023). Telecommuting-related health outcomes during the COVID-19 pandemic in South Korea: A national population-based cross-sectional study. *BMC Public Health*, *23*(1), 1–11. https://doi.org/10.1186/s12889-023-15271-0

Ryoo, S.-W., Min, J.-Y., Son, S.-Y., Choi, B.-Y., Choi, J., & Min, K.-B. (2024). Telecommuting during the COVID-19: The role of job demand and control on health outcomes. *Annals of Occupational and Environmental Medicine*, *36*(1), 1–13. https://doi.org/10.35371/aoem.2024.36.e14

Schmitz, H., Bauer, J. F., & Niehaus, M. (2023). Working anytime and anywhere-even when feeling ill? A cross-sectional study on presenteeism in remote work. *Safety and Health at Work*, *14*(4), 375–383. https://doi.org/10.1016/j.shaw.2023.11.001

Shafer, L., Ahmed, F., Kim, S., Wernli, K. J., Jackson, M. L., Nowalk, M. P., Bear, T., Zimmerman, R. K., Martin, E. T., Monto, A. S., Gaglani, M., Reis, M., Chung, J. R., Flannery, B., & Uzicanin, A. (2023). Relationship between telework experience and presenteeism during COVID-19 pandemic, United States, March–November 2020. *Emerging Infectious Diseases*, *29*(2), 278–285. https://doi.org/10.3201/eid2902.221014

Spinks, W. A. (2002). A survey of home-based workers in Japan: Emerging health issues. *Journal of Occupational Health*, *44*(4), 248–253. https://doi.org/10.1539/joh.44.248

Steidelmüller, C., Meyer, S.-C., & Müller, G. (2020). Home-based telework and presenteeism across Europe. *Journal of Occupational and Environmental Medicine*, *62*(12), 998–1005. https://doi.org/10.1097/JOM.0000000000001992

Wang, B., Liu, Y., Qian, J., & Parker, S. K. (2021). Achieving effective remote working during the COVID-19 pandemic: A work design perspective. *Applied Psychology*, *70*(1), 16–59. https://doi.org/10.1111/apps.12290

Zhao, D., & Strotmann, A. (2015). Analysis and visualization of citation networks. In *Synthesis Lectures on Information Concepts, Retrieval, and Services*. Springer International Publishing. https://doi.org/10.2200/S00624ED1V01Y201501ICR039

Exploring the relationship between spatial factors and poverty in Indonesia based on macroeconomics: A look at Java Island

***Thomas Ola Langoday*[1]**

[1] *Sekolah Tinggi Ilmu Ekonomi (Stie) Oemathonis Kupang, Indonesia, ORCID: 0009-0006-0984-0876, thomasolalangoday01@gmail.com (corresponding author).*

Abstract: *Poverty is not only a problem within a region but also a problem that affects multiple regions. The current efforts to alleviate poverty have primarily focused on the household and community level, often neglecting the crucial role that spatial dimensions play in understanding poverty dynamics. The theory of the vicious circle of poverty still needs attention from the whole community and the government, because Indonesia is committed to alleviating poverty. This study examines the significance of spatial considerations in poverty alleviation efforts across 117 districts on Java Island. The research aims to provide recommendations for poverty alleviation policies considering spatial dimensions, including regional conditions and interactions. It was found that almost all macroeconomic variables are spatially dependent, except inflation and health spending. The Moran I measure revealed that the spatial correlation was considerable. It was determined by utilizing spatial econometric techniques. Specifically, with the spatial autoregressive and spatial Durbin model methods, poverty incidence on Java Island highly depends on spatial factors. The study indicated that investing in education in neighboring areas and growing industrial sectors considerably lower poverty within a specific district. The report recommends that policies be executed in a coordinated manner to effectively reduce poverty on Java Island, focusing on industrial and human development through education spending. The spatial network parameters indicate that the impact of these variables is still relatively small, but the effect is specific and accurate. Based on the study's results, several suggestions for addressing poverty in Java are provided. The government should improve the connections between different areas (provinces, districts, and cities) to increase access among regions.*

Keywords: *Poverty, spatial dependence, regional management, spatial-spillover effects, poverty management strategies.*

JEL Classification: *I32, O18, R11, R58.*

APA Style Citation: Langoday, T. O. (2026). Exploring the relationship between spatial factors and poverty in Indonesia based on macroeconomics: A look at Java Island. *E&M Economics and Management, 29*(2), 32–44. https://doi.org/10.15240/tul/001/2026-2-003

Introduction

Compared to other ASEAN nations, which have rates of less than 5%, Indonesia and the Philippines have had relatively high rates of poverty incidence over the past ten years, ranging from 6 to 9% annually (Miar & Yunani, 2020). Economic theory and social interaction-based economic and social variables have been the main research topics on poverty in these nations. It is crucial to remember that Indonesia and the Philippines are archipelagos with distinctive characteristics that may impact poverty differently than in other ASEAN countries (Aji, 2022).

Indonesia, specifically, has regional economic disparities, leading to distinct characteristics of poverty in each region (Hidayah et al., 2023; Miar & Yunani, 2020). Generally, poverty in Java is lower than in other islands, but in terms of numbers, there are more poor people in Java than in other islands. As a result, poverty alleviation strategies in Indonesia are often implemented at the macro and micro levels but rarely consider spatial factors (Ardi & Isnayanti, 2020; Suryahadi et al., 2020).

Several studies have found that effective poverty reduction policies in Indonesia include macroeconomic policies, targeted social assistance, and spatial policies (Hasan, 2021; Primanto et al., 2021). However, spatial aspects should be more noticed in poverty alleviation strategies, which can hinder their effectiveness in regions with diverse poverty characteristics (Aji, 2022; Miar & Yunani, 2020).

This research is based on the vicious cycle of poverty theory. The theory states that the occurrence of poverty is caused by human backwardness and natural resources (Addae-Korankye, 2019; De Bruijn & Antonides, 2022). which means that poverty is not caused only by the community itself, but has links with other resources that influence each other (Brady, 2019; Tehubijuluw et al., 2021). In this study, the theory of the vicious circle of poverty still needs attention from the whole community and the government, because Indonesia is committed to alleviating poverty (De Bruijn & Antonides, 2022; Soegoto et al., 2022). Proven from year to year, the number of poverties can be reduced. This is also proven macro-economically that the Indonesian people are able to rise from poverty through various strategies offered by the government, namely SMEs, agriculture, human resource capacity building and a number of other poverty management strategies (Brady, 2019; Tehubijuluw et al., 2021).

Given the importance of spatial factors in poverty alleviation, this study aims to investigate the significance of spatial aspects in reducing poverty in Indonesia, particularly in Java (Purwono et al., 2021; Suryahadi et al., 2020). The policy contribution of this research is to provide support for making better regulations to increase government and community efforts to continue to eradicate poverty (Hasan, 2021). Empirical contributions are directed to the concept of poverty alleviation strategies, and improvement of access to consumption of social services (education, health and nutrition) (Suryahadi et al., 2020). Expanding the scope and quality of basic services requires capital investment, which will ultimately increase the productivity of the poor. At the same time, these services will satisfy the basic consumption needs of the poor. Basic services such as clean water, garbage disposal, housing and others (Ardi & Isnayanti, 2020). Furthermore, the theoretical contribution is aimed at the circle of poverty theory, which states that poverty is caused by human backwardness and natural resources (Hidayah et al., 2024). Management of natural resources is highly dependent on human productive capacity (Soegoto et al., 2022; Tehubijuluw et al., 2021). All transactions in banking can now be accessed via Internet banking without being limited by place or time; in other words, through Internet banking, any form of transaction can be done quickly and easily. Internet banking has become a significant concern and a revolutionary strategic weapon for bank operations, delivery, and competition between banks (Candrawati & Widiastuti, 2024). If the population is poor and their education is low, it will result in a scarcity of technical skills, knowledge and entrepreneurial activities, which will automatically cause the available natural resources to be neglected, underdeveloped and even misused (Miar & Yunani, 2020; Purwono et al., 2021).

1 Theoretical background

1.1 Vicius circle of poverty theory

The occurrence of poverty is caused by backwardness of human and natural resources. Management of natural resources is highly dependent on human productive capacity (Priyangga & Yasyfi, 2020). If the population is poor and their education is low, it will result in a scarcity of technical skills, knowledge and entrepreneurial activities, which will automatically cause the available natural resources to be neglected, underdeveloped and even misused (Konovalova, 2022). On the other hand, the lack of natural resources will lead to poverty, because natural resources are the main source of needs in human life (Dubovik, 2022). Poverty of natural resources is both a cause and a result of human poverty (Pang et al., 2022). The agricultural and human development sectors play an important role in the economic development of poverty reduction in Indonesia (Elisha & Felix, 2021). Aspects of human and agricultural development have contributed significantly to poverty reduction, especially in rural areas. The biggest contribution to increasing rural income and reducing poverty in rural areas resulted from the technological revolution in rice farming, including irrigation development (Cheng et al., 2019; Dubovik, 2022; Konovalova, 2022).

1.2 Poverty in Indonesia

Studies on poverty subjects from 2008 to 2021 can be plotted using a bibliometric method based on the ScienceDirect database. Early research on poverty concentrated on rural poverty (Alamanda, 2020). Still, by 2008, the focus had switched to urban poverty, according to an examination of the network of themes connected to poverty from earlier studies (Alamanda, 2020; Sambodo & Novandra, 2019). Around 2012, studies expanded to cover more significant issues like poverty alleviation or reduction (Santika et al., 2021). The focus turned to fuel poverty, spatial variance, and neighborhood poverty in 2014. Recent studies have mainly dealt with the spatial analysis of poverty, emphasizing energy poverty (Zhu et al., 2021). Asset management is an essential element that forms the basis for preparing regional government balance sheets. Therefore, its management must be carried out in an orderly and systematic manner (Mokoginta et al., 2024).

Spatial analysis is similar to regional analysis, but it is distinguished by its focus on specific coordinates and their interactions and relationships. It also ranges from physical distance to economic and social distance (Haini, 2020). In Indonesia, regional analysis plays a crucial role in understanding the dynamics of various regional development issues due to differences in geographical conditions, natural resources, demographics, and regional autonomy systems. However, it is distinct from spatial analysis (Li et al., 2022; Zhu et al., 2021).

Development issues such as poverty must be approached with a more comprehensive perspective. Qin and Zhang (2022) classify the root causes of poverty into two factors: economic and spatial. Compensation is all income from money, goods, or indirect income that employees receive in return for services rendered. Usually, employees will cheat because of dissatisfaction or disappointment with the results of the compensation they receive for what has been done (Dewi et al., 2024). Economic factors are related to household and community capacity to access economic opportunities and improve welfare, including improvements in education and health, changes in employment type and status, reductions in the number of dependents, and increases in asset ownership (Li et al., 2022; Sambodo & Novandra, 2019).

In contrast to economic factors, spatial factors can significantly impact poverty due to the condition and physical position of regions (Alamanda, 2020). The fixed position of a region can provide opportunities for basic infrastructure, management of natural resources, and critical land to be developed, which can mitigate the effects of natural disasters in isolated areas (Hasan, 2021; Miar & Yunani, 2020). However, limited resources and geographical barriers can also create natural poverty, which leads to low productivity among local populations. The position of a region also determines the degree of access, level of interrelation, and interaction among regions. Physical distance is a primary factor in spatial analysis, with longer distances leading to higher costs in terms of time and money (Aji, 2022; Ardi & Isnayanti, 2020; Sambodo & Novandra, 2019).

Furthermore, interaction with neighboring regions is believed to contribute to poverty. Some researchers argue that regional growth is not only dependent on the ability of regions to increase their production activities and labor and capital mobility among regions in terms of differences in wages and returns on regional investment (Haini, 2020; Santika et al., 2021; Zhu et al., 2021). Transportation is the main instrument used to facilitate mobility among regions (Sambodo & Novandra, 2019). Many studies have investigated the impact of transportation on spatial analysis. Zhu et al. (2021) argue that transportation enables the mobility of resources among regions, creating economic value for growth and welfare. Li et al. (2022) add an interregional income model, where the role of government, both in terms of income and expenditure, is considered an endogenous factor that fluctuates based on trade activities among regions. Poverty among regions can occur when the mobility of these elements is hindered in certain areas, thereby threatening the surrounding area (Qin & Zhang, 2022).

Regions that undertake the development of a base sector can positively and negatively impact the surrounding area (Z. Wang et al., 2023). Positive effects include the development and creation of new jobs through the transfer of physical resources and innovation and knowledge. Economic expansion positively and significantly reduces poverty. It means that if economic growth increases, poverty also increases (Jayawarsa et al., 2023). However, some areas may lose potential resources and investment, which can cumulatively make these areas poorer. Thus, spatial interaction can be either mutualistic or parasitic (Xu et al., 2022).

2 Research methodology

This study utilized data collected from the Central Bureau of Statistics (BPS) to analyze the impact of spatial factors on poverty decline in the districts of Java (Addae-Korankye, 2019; Brady, 2019). The data covered 2010 to 2022 and included 117 districts except Kepulauan Seribu. In order to determine the influence of spatial factors on poverty reduction in these districts, spatial statistics were applied, precisely Moran's I test on response and predictor variables (Aji, 2022; Miar & Yunani, 2020).

The study followed a specific process to conduct this analysis. Firstly, the model was constructed based on the theoretical foundation, which involved a comprehensive review of relevant theories and literature to identify key variables essential for understanding poverty dynamics within Java's districts. The theoretical foundation guided the selection of appropriate variables such as economic growth ($ECO = X_1$), industry sector ($IND = X_2$), inflation ($INFL = X_3$), education spending ($EDU = X_4$), health spending ($HEA = X_5$), and infrastructure spending ($IFST = X_6$), which are known to have a significant impact on poverty alleviation ($POV = Y$). Thus, the basic model was formulated as follows.

$$POV_{it} = \beta_{0i} + \beta_{1i}ECO_{it} + \beta_{2i}IND_{it} + \\ + \beta_{3i}INFL_{it} + \beta_{4i}EDU_{it} + \beta_{5i}HEA_{it} + \\ + \beta_{6i}IFST_{it} + E_{it} \tag{1}$$

Which can be explained:

$$Y_{it} = \beta_{0i} + \beta_{1i}X_{1it} + \beta_{2i}X_{2it} + \beta_{3i}X_{3it} + \\ + \beta_{4i}X_{4it} + \beta_{5i}X_{5it} + \beta_{6i}X_{6it} + E_{it} \tag{2}$$

Or in matrix form:

$$P_t = A + X_tB + E_t \tag{3}$$

$$P_t = \begin{bmatrix} P_{1t} \\ P_{2t} \\ \vdots \\ P_{Kt} \end{bmatrix};\ A = \begin{bmatrix} \alpha_{1t} \\ \alpha_{2t} \\ \vdots \\ \alpha_{Kt} \end{bmatrix};\ X = \begin{bmatrix} X_{11t} & X_{12t} & \cdots & X_{K1} \\ X_{21t} & X_{22t} & \cdots & X_{K2} \\ \vdots & \vdots & \ddots & \vdots \\ X_{K1t} & X_{K2t} & \cdots & X_{KL} \end{bmatrix}; \tag{4a}$$

$$B = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1L} \\ b_{21} & b_{22} & \cdots & b_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ b_{K1} & b_{K2} & \cdots & b_{KL} \end{bmatrix};\ E = \begin{bmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{Kt} \end{bmatrix} \tag{4b}$$

Secondly, a crucial step involved the creation of a spatial weight matrix, which served as a fundamental component in the spatial analysis, enabling the quantification of spatial relationships between different geographic units. Several alternatives for constructing the weight matrix were considered, including exploring different distance metrics, neighbor definitions, and weighting schemes to ensure the most appropriate representation of spatial relationships within the analysis. In the second step, the author builds the spatial weighting (*W*) by applying relative to total transport costs among districts. The higher the relative cost, the lower the spatial weight. The complete spatial weight can be written in matrix *W* as follows.

$$W = \begin{bmatrix} W_{11} & W_{12} & \cdots & W_{1K} \\ W_{21} & W_{22} & \cdots & W_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ W_{K1} & W_{K2} & \cdots & W_{KK} \end{bmatrix} \tag{5}$$

where: W – the spatial weight for a district to district; i and j – district identity; K = 177 (the number of districts used in this study); L = 6 (the number of independent variables).

The third step is performed to build specific spatial econometric panel data. Four models can be investigated.

1. **Spatial lag model (SAR)**
$$P_t = A + \lambda WP_t + X_tB + E_t \tag{6}$$
2. **Spatial error model (SEM)**
$$P_t = A + X_tB + E_t \\ E_t = \lambda WE_{t-1} + \upsilon_t \tag{7}$$
3. **Spatial Durbin model (SDM)**
$$P_t = A + \lambda WP_{t-1} + WX_tB + E_t \tag{8}$$
4. **Spatial Durbin error model (SDEM)**
$$P_t = A + \lambda WP_{t-1} + WX_tB + E_t \\ E_t = \lambda WE_{t-1} + \upsilon_t \tag{9}$$

Next, a spatial-specific model was decided. This decision was informed by the unique characteristics of the dataset and the research objectives, aiming to capture the spatial dependencies and autocorrelation inherent in the poverty data of Java's districts. Not necessarily that all models have to bring about the same results. Choosing the best or appropriate model is the most important step in investigating the magnitude of spatial spillover from various regional conditions, such as regional economic concentration and elements of government policy.

Finally, the model was computed using data collected and analyzed with R software, a powerful programming language and environment for statistical computing and graphics. Utilizing R software enabled efficient processing of the collected data, implementation of the chosen spatial model, and derivation of meaningful insights into the spatial patterns and drivers of poverty decline within Java's districts. Panel data testing was performed using the Chow and Hausman tests to characterize the analyzed data (Saputra et al., 2023).

3 Results and discussion

3.1 Poverty line in Java Island, Indonesia

The Central Statistics Agency records that there are around 26.5 million poor people in Indonesia. The province on Java Island that has the highest poverty line per capita in September 2021 is DKI Jakarta, which is IDR 715,052 (45 USD/capita/month). After that, there is Banten with a poverty line of IDR 547,483 (34 USD/capita/month) and DI Yogyakarta IDR 496,904 (31 USD/capita/month). Meanwhile, the poverty line in East Java is IDR 445,139 (28 USD/capita/month), West Java is IDR 437,604 (27 USD/capita/month), and Central Java is IDR 423,264 (26 USD/capita/month) according to Primanto et al. (2021), Purwono et al. (2021) and Soegoto et al. (2022). This statement is shown in Fig. 1.

The poverty line is the minimum income level to obtain a decent standard of living in an area. Residents with an average expenditure below the poverty line are categorized as poor (Ardi & Isnayanti, 2020; Haini, 2020; Hasan, 2021).

3.2 Spatial dependence test result

The data provided summarizes the results of a spatial dependence test, specifically a Moran's I test. The test measures the spatial autocorrelation of a response variable (in this case, poverty) and its predictor variables (human capital index, economic growth, industry sector, education spending, health spending, and infrastructure spending). This result is exactly what China experienced. China is one of the countries with the highest levels of poverty and hunger in the world. High poverty and hunger in China are caused by several things, including large amounts of urbanization, lack

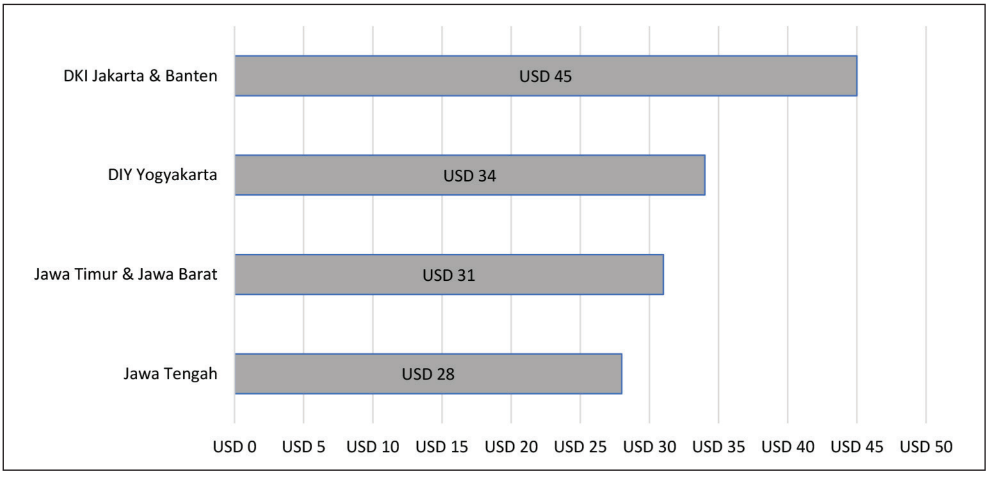

Fig. 1: Graph of poverty in Java

Source: own based on OECD (2021)

of education, difficulty accessing healthcare, agricultural lifestyle, Hukou System and migrant workers (Li et al., 2022; Qin & Zhang, 2022).

The findings show a positive spatial dependency between regions, which means that the values of the response and predictor variables positively correlated with those of their surrounding areas (Z. Wang et al., 2023; Zhu et al., 2021). Although there is an opportunity for development in regional integration, the correlation density is modest. If we compare these findings in Indonesia with conditions in other countries such as China, this condition is like what happened in China, experiencing poverty and hunger, a low human development index and a level of health that has not yet reached rural areas. Based on China's geographical information system (GIS), poverty and hunger in China mostly occur in rural areas. There are 26 million people living in extreme poverty (Li et al., 2022; Xu et al., 2022).

Tab. 1 shows the values and probabilities of the global Moran I test and the Moran test for poverty and each predictor variable. The values of Moran's I range from 0.16 to 0.25, and the probabilities range from 0.01 to 0.001, which indicates that the results are statistically significant at a significance level of 0.001 (Saputra, 2021).

It should be noted that these results were obtained using RStudio version 1.2.1335 (2019) and processed by the author. It is essential to consider any limitations or assumptions made in the study in interpreting the results.

Tab. 1: Spatial autocorrelation

	Value	Probability
Global Moran 1	0.24	0.00
Moran test Poverty (*Y*)	0.25	0.00
Moran test predictors variable		
Humam capital index (*HDI*)	0.16	0.00
Economy growth (*ECO*)	0.07	0.01
Industry sector (*IND*)	0.02	0.05
Education spending (*EDU*)	0.18	0.00
Health spending (*HEA*)	0.01	0.24
Infrastructure spending (*INFL*)	0.20	0.00

Source: own processed in Software RStudio version 1.2.1335 (2019)

3.3 Results of estimated spatial models

The data provided results from a spatial dependence test, which measures the spatial autocorrelations of the response variable (poverty) and its predictor variables (human development index, economic growth, industry sector, education spending, health spending, and infrastructure spending) (Setiawan, 2020; Q. Wang et al., 2023). The test findings indicate that there is a beneficial impact from one region to another but that there is potential for development in regional integration because the correlation density is relatively low (Xu et al., 2022). The data also shows the results of some tested models, including the spatial autoregressive (SAR) and spatial durbin model (SDM) (Setiawan, 2020). The models were compared to a fixed effect model with dummy provinces to determine which provinces had the most effective efforts in reducing poverty during the study period (Sucipto et al., 2022).

With high adjusted R-squared values of 0.98, the SAR and SDM models successfully account for a sizable portion of the variation in poverty. The intercept values for both models are also significant (at the 0.001 level). The spatial autoregressive coefficient (ρ) for both models is also significant (at the 0.001 level), indicating a strong spatial dependence in the data (Luwihono et al., 2021; Q. Wang et al., 2023). The coefficients for the original variables (human development index, economic growth, industry sector, education spending, health spending, and infrastructure spending) show that human development index, education spending, and industry sector have a negative and significant relationship with poverty (at the 0.001 level) (Sara et al., 2023). In contrast, economic growth, health spending, and infrastructure spending have a negative but not significant relationship with poverty (Saputra et al., 2023).

The coefficients for the spatially lagged variables (Lag_of industry sector and Lag_of education spending) show that they have a negative and significant relationship with poverty (at the 0.05 level) (Setiawan, 2020). The coefficients for the dummy provinces (D1_Banten, D2_DKI, D3_WestJava, D4_Central Java, D5_EastJava) show that all provinces have a negative and significant relationship with poverty (at the 0.001 level). It implies that these provinces have made practical efforts to reduce poverty during the study period.

These results are in line with panel research in the European Union. Bosco and Poggi (2020) analyzed the relation between poverty dynamics and observable and unobservable country factors using longitudinal 2008–2011 EU-SILC data for 26 European countries. They estimated a three-level dynamic model. The three levels considered were individual, time, and country. They included micro-level determinants of poverty, country-level variables, lagged poverty, and initial conditions. They found evidence

Tab. 2: Estimations outputs

	FE with dummy		SAR		SDM	
	Value	Prob.	Value	Prob.	Value	Prob.
Adjusted R^2	0.65		0.98		0.98	
Intercept	62.93	0.00	22.86	0.00	26.73	0.05
Spatial autoregressive			0.61	0.00	0.57	0.04
Original variables						
Human development index	−0.60	0.00	−0.24	0.00	−0.27	0.05
Economic growth	−0.04	0.05	−0.04	0.04	−0.04	0.02
Industry sector	−0.04	0.00	−0.02	0.03	−0.02	0.05
Education spending	−1.31	0.00	−0.63	0.05	−0.19	0.54
Health spending	−1.16	0.17	−0.60	0.39	−0.69	0.33
Infrastructure spending	−0.93	0.05	−0.00	0.99	0.42	0.36
Spatially lagged variables						
Lag_of industry sector					−0.05	0.05
Lag_of education spending					−1.15	0.04
Dummy province						
D1_Banten	−13.10	0.01				
D2_DKI	−9.23	0.00				
D3_WestJava	−9.34	0.00				
D4_CentralJava	−5.20	0.01				
D5_EastJava	−7.10	0.00				

Source: own processed on Software RStudio version 1.2.1335 (2019) of genuine state dependence, as well as a role of the initial value of poverty. They also found important evidence of unobserved heterogeneity across individuals.

Overall, the data suggest a strong spatial dependence on the poverty data and that certain variables, such as the human development index, education spending, the industry sector, and certain provinces, significantly reduce poverty in the study region (Sambodo & Novandra, 2019; Santika et al., 2021).

3.4 Discussion

The estimation results of this study indicate that spatial factors significantly impact poverty on Java Island, with autoregressive spatial coefficients ranging from 0.57 to 0.61. Despite the relatively low coefficients, the effect is deemed to be accurate (Haini, 2020; Setiawan, 2020). This leads the author to conclude that the impact of surrounding areas on a district/city is indeed present. However, the strength of the relationship is relatively tiny. If access between areas in the system were to be improved, the impact size would likely be greater (Luwihono et al., 2021; Sucipto et al., 2022).

Except for the health spending variable, which is found to have no significant impact in explaining poverty reduction, the study's findings also show that all variables in the models have the predicted signals. The signs of coefficients for all variables are consistent across all specifications (Hafizh et al., 2020). The models consistently illustrate that the quality of human resources plays an essential role in reducing poverty, followed by the proportion of education spending and industrial activities as the basis of the economy (Q. Wang et al., 2023; Z. Wang et al., 2023).

Based on research findings, it is similar to what happened in the European Union. Giarda and Moroni (2018) exploited the longitudinal component of the 2009–2012 EU-SILC data for France, Italy, Spain, and the UK to estimate the degree of poverty state dependence. They estimated different specifications of a dynamic random effects probit model to disentangle the role of regional disparities within countries. Their findings suggest that there is evidence of genuine state dependence in all of those countries. In comparative terms, when not accounting for regional disparities within countries the degree of poverty persistence is highest in Italy and lowest in the UK. When regional effects are included, the degree of poverty persistence in Italy drops, suggesting that unlike other countries, in Italy regional disparities play an important role in explaining poverty state dependence. The conditions that occur in the European Union are also caused by and have an impact on local government spending, so that handling is relatively slow and requires adequate solutions. The development of the banking industry has undergone significant changes in recent years, especially in the banking world in Indonesia, which is increasingly crowded with Islamic banks, where operational activities are different from conventional banks (Badawi et al., 2023).

Local government spending, such as spending on education and infrastructure, has various levels of impact on poverty. In Java, education spending is focused on efforts to achieve the success of the twelve-year compulsory education program and non-formal education programs (Ardi & Isnayanti, 2020). Research has shown that local government budget policies in the education sector are effective attempts at investing in human capital in developing regions (Pang et al., 2022). Through budget allocations in education programs, an increase in the quality and accessibility of society towards education is expected (Hasan, 2021). Several studies have shown that government spending in various sectors positively impacts poverty reduction (Cheng et al., 2019). The World Bank found that spending on infrastructure, education and health, social protection, agriculture, rural development, and employment programs, respectively, are effective in reducing poverty (Alamanda, 2020). These findings support the conclusion that local government spending significantly impacts poverty alleviation in Indonesia (Elisha & Felix, 2021). However, there are some notes derived from studies by the World Bank that government spending on corruption, ineffective programs, and military and defense, respectively, can have a negative impact on poverty reduction (Haini, 2020; Santika et al., 2021). This raises concerns about the effectiveness of local government spending in Indonesia, where corruption and mismanagement of funds are prevalent issues (Dubovik, 2022; Pang et al., 2022; Zhou & Huang, 2023). In line with these findings, regions with a large allocation of education spending tend to have an elevating human development index (Li et al., 2022). The aim of the company selling shares in the capital market is not only to achieve an optimal financial structure but what is more importantly, to increase the value of the company with the hope that the company will be able to provide welfare to its investors, managers, and employees so that the company can raise a positive image towards its customers (Arisudhana, 2023).

In this study, spending on infrastructure was found to be significant only in the fixed effect model with a dummy province, measured by spending on public works, planning, and transportation (Qin & Zhang, 2022). Infrastructure expenditure aims to improve facilities and infrastructure supporting the economic sector (Xu et al., 2022). The allocation of infrastructure spending in Java is expected to reduce poverty through the construction of facilities and infrastructure in the form of tolls, electricity, and clean water to increase the productivity of the interconnected areas (Sucipto et al., 2022). According to previous research, infrastructure spending can impact spillover by creating labor and additional investments to reduce transportation costs (Luwihono et al., 2021). Furthermore, connectivity among regions through infrastructure development encourages economic opportunities for the surrounding area and further reduces poverty (Hafizh et al., 2020).

The study also used an SDM model to obtain the significance of the direct and indirect impacts of the independent variables, which indicates that there is spatial interaction of the independent variables in influencing the poverty level of districts/cities in Java (Sameti et al., 2022). The direct effect is depicted by an impact directly occurring in the area. In contrast, the indirect effect includes the spatial spillover or the interaction of the area with the region (Setiawan, 2020; Sucipto et al., 2022). The author found a positive spillover effect in poverty reduction through education spending and industrial activity at the local level (Hafizh et al., 2020; Luwihono et al., 2021). This finding is also applicable in regions with an industrial base with spillover effects on poverty alleviation compared to the non-industrial base regions (Olilingo & Putra, 2020).

Research by Stoeffler and Joseph (2020) points out that in the long run, education spending creates a positive externality for the surrounding area. Regions with more lavish education spending can attract workers with high skills and knowledge spillovers. Furthermore, it also encourages employment growth. Sameti et al. (2022) suggest that policymakers should pay much attention to their neighbors before deciding how much to spend on public services and accelerating the industry sector. The study found that education spending has a significant impact on poverty alleviation in Indonesia, which is consistent with other studies that emphasized aspects of human capital, access to employment, and improved health outcomes (Olilingo & Putra, 2020).

The benefits of agglomeration, however, can also exacerbate poverty through the polarization of regional growth centers. In these instances, surrounding areas cannot take advantage of the demand generated by industrial areas, leading to a decline in productivity and a lack of investment and capital for regional development (Li et al., 2022; Stoeffler & Joseph, 2020; Sucipto et al., 2022). This phenomenon is observed in East Java, where the development of the industrial sector tends to have a more significant negative impact on surrounding areas than a positive spread effect. The study found a tendency for investment flows to concentrate in the Surabaya, Sidoarjo, Pasuruan, and Mojokerto regions, as indicated by an increase in the minimum wage rate. As a result, surrounding areas experience less productivity and fewer improvements in basic infrastructure, leading to increased poverty (Hafizh et al., 2020; Luwihono et al., 2021; Sucipto et al., 2022). The benefits of agglomeration, such as improved economic efficiency, access to employment, increased labor productivity and wages, and increased provision of public goods and services, can lead to poverty reduction. However, inadequate allocation of resources to rural areas, poor-quality public goods and services, increased inequality between rich and poor, poor infrastructure, and inflation can limit the benefits of agglomeration in reducing poverty (Olilingo & Putra, 2020; Stoeffler & Joseph, 2020; Z. Wang et al., 2023).

Furthermore, this finding is supported by the significance of the local dummy variable, which highlights the differences in the effectiveness of inter-provincial poverty reduction policies (Sameti et al., 2022; Setiawan, 2020). Provinces such as Banten, DKI, and West Java have a more significant potential to quickly reduce poverty among regencies and cities, thanks to the positive impacts of industrial development and the advantages of agglomeration (Hafizh et al., 2020; Olilingo & Putra, 2020). Precisely, industrial development in these provinces results in increased large-scale demand, the growth of new industries, and improved infrastructure in the form of adequate water and electricity facilities. The cumulative process of these factors helps to encourage the spread of effects and reduce barriers to development, ultimately leading to a decrease in poverty, particularly in industry-based areas (Sameti et al., 2022; Stoeffler & Joseph, 2020; Sucipto et al., 2022).

Overall, the general characteristic of the studies reviewed is that they find evidence of high rates of poverty that still exist, especially in regions of Indonesia in direct proportion to conditions in Southern European countries and China, as well as the role of unobserved heterogeneity. However, these studies do not investigate the evolution of poverty components over time. Given the importance of the phenomenon of poverty and its heterogeneity in Indonesia today, there is an important need to study its determinants and evolution over time.

Conclusions

Poverty is not only a problem within a region but also a problem that affects multiple regions. This study has aimed to explore the significance of spatial factors in poverty reduction efforts in Indonesia, particularly focusing on the region of Java. By addressing this objective, the research contributes to the formulation of more effective poverty alleviation strategies, emphasizing the importance of spatial considerations in policymaking. This study provides evidence that poverty is a spatially dependent issue on the island of Java. The analysis of macroeconomic variables included in the model shows that changes in these variables in one region can impact poverty incidence in surrounding areas (He, 2019). Additionally, the study suggests that macroeconomic variables that significantly impact poverty in a specific region are the development of industrial sectors and spending on education. Increased development of industrial sectors in surrounding areas of a district can lead to a significant reduction in poverty within that district (Zhou & Huang, 2023). The same holds for spending on education.

The spatial network parameters indicate that the impact of these variables is still relatively small, but the effect is specific and accurate (Priyangga & Yasyfi, 2020). Based on the study's results, several suggestions for addressing poverty in Java are provided. The government should improve the connections between different areas (provinces, districts, and cities) to increase access among regions (Konovalova, 2022). The stronger the connections between regions, the more significant the impact on poverty (Dubovik, 2022). Policy decisions should prioritize industrial development and spending on education as significant spatial impact variables. Public policy should achieve synergy between industrial and human development through education spending.

While the findings of the study provide valuable insights into the spatial dynamics of poverty and the effectiveness of poverty reduction efforts in Java Island, Indonesia, it is important to interpret them with caution due to potential limitations. Unaccounted-for variables or spatial patterns could influence the observed relationships between poverty and its predictors. The spatial models may not have captured all relevant factors shaping poverty dynamics in Java Island, such as cultural, historical, or institutional differences across regions (Sugiharti et al., 2022). Additionally, while significant relationships were found between certain predictor variables and poverty, correlation does not imply causation. The study's focus on Java Island may limit its generalizability to the broader context of poverty dynamics in Indonesia, as different regions may face unique challenges. Moreover, reliance on secondary data sources introduces the possibility of limitations such as data quality and consistency over time. To improve future research, it is important to include a broader range of variables and consider regional differences within Indonesia. Using advanced methods, ensuring data quality, and engaging stakeholders can enhance the reliability and applicability of findings for informed poverty reduction efforts.

Acknowledgments: *The authors would like to thank all parties who supported this research. Professional support and assistance from all respected reviewers have made this article qualified to be published.*

References

Addae-Korankye, A. (2019). Theories of poverty: A critical review. *Journal of Poverty,* *Investment and Development*, *48*(1), 55–62. https://doi.org/10.7176/JPID/48-08

Aji, Y. I. T. (2022). The effect of village income and gross regional domestic product on poverty in Indonesia. *Jurnal Bina Praja. Journal of Home Affairs Governance*, *14*(2), 315–328. https://doi.org/10.21787/jbp.14.2022.315-328

Alamanda, A. (2020). The effect of government expenditure on income inequality and poverty in Indonesia. *Info Artha*, *4*(1), 1–11. https://doi.org/10.31092/jia.v4i1.614

Ardi, N. & Isnayanti. (2020). Structural equation modelling-partial least square to determine the correlation of factors affecting poverty in Indonesian provinces. *IOP Conference Series: Materials Science and Engineering*, *846*(1), 012054. https://doi.org/10.1088/1757-899x/846/1/012054

Arisudhana, D. (2023). The influence of some financial ratios, sales growth and debt maturity on financial performance. *Journal of Governance, Taxation and Auditing*, *2*(2), 123–132. https://doi.org/10.38142/jogta.v2i2.915

Badawi, A., Nugroho, L., & Hidayah, N. (2023). Determinant factors of Islamic bank financial performance and competitive advantage as moderating variables in Islamic banks in Indonesia. *International Journal of Environmental, Sustainability, and Social Science*, *4*(1), 37–44. https://doi.org/10.38142/ijesss.v4i1.270

Bosco, B., & Poggi, A. (2020). Middle class, government effectiveness and poverty in the EU: A dynamic multilevel analysis. *Review of Income and Wealth*, *66*(1), 94–125. https://doi.org/10.1111/roiw.12403

Brady, D. (2019). Theories of the causes of poverty. *Annual Review of Sociology*, *45*(1), 155–175. https://doi.org/10.1146/annurev-soc-073018-022550

Candrawati, T., & Widiastuti, R. (2024). The influence of perceived usefulness, perceived ease of use, and perceived risk on the level of customer satisfaction in using internet banking (Study on deposit customers of Bank Syariah Indonesia Kcp Malang Kawi). *International Journal of Environmental, Sustainability and Social Science*, *5*(1), 139–146. https://doi.org/10.38142/ijesss.v5i1.984

Cheng, H., Dong, S., Li, F., Yang, Y., Li, Y., & Li, Z. (2019). A circular economy system for breaking the development dilemma of 'ecological Fragility – Economic poverty' vicious circle: A CEEPS-SD analysis. *Journal of Cleaner Production*, *212*, 381–392. https://doi.org/10.1016/j.jclepro.2018.12.014

De Bruijn, E.-J., & Antonides, G. (2022). Poverty and economic decision making: A review of scarcity theory. *Theory and Decision*, *92*(1), 5–37. https://doi.org/10.1007/s11238-021-09802-7

Dewi, R. O. C., Saputra, K. A. K., & Pramartha, I. M. A. (2024). The effect of internal control system, information asymmetry, and suitability of compensation on the tendency of accounting fraud (fraud) in Denpasar City (Opd). *Journal of Governance, Taxation and Auditing*, *2*(4), 246–252. https://doi.org/10.38142/jogta.v2i4.1014

Dubovik, M. V. (2022). The digital economy – The service economy for the poor or the poverty digitalization. In *Business 4.0 as a subject of the digital economy* (pp. 1119–1124). Springer International Publishing. https://doi.org/10.1007/978-3-030-90324-4_185

Elisha, O. D., & Felix, M. J. (2021). Destruction of coastal ecosystems and the vicious cycle of poverty in Niger Delta Region. *Journal of Global Agriculture and Ecology*, *11*(2), 7–24.

Giarda, E., & Moroni, G. (2018). The degree of poverty persistence and the role of regional disparities in Italy in comparison with France, Spain and the UK. *Social Indicators Research*, *136*(1), 163–202. https://doi.org/10.1007/s11205-16-1547-3

Hafizh, M., Hidayah, N., & Silalahi, P. R. (2020). Macroeconomics and profit-sharing financing in Islamic banking in Indonesia: The third parties fund as intervening. *Jurnal Akuntansi Dan Keuangan Islam*, *8*(2), 131–147. https://doi.org/10.35836/jakis.v8i2.183

Haini, H. (2020). Spatial spillover effects of public health and education expenditures on economic growth: Evidence from China's provinces. *Post-Communist Economies*, *32*(8), 1111–1128. https://doi.org/10.1080/14631377.2020.1722586

Hasan, Z. (2021). The effect of economic growth and human development index on poverty in Indonesia. *Journal of Economics and Sustainability*, *3*(1), 42–53. https://doi.org/10.32890/jes2021.3.1.5

He, X. (2019). Digital entrepreneurship solution to rural poverty: Theory, practice and policy implications. *Journal of Developmental Entrepreneurship*, *24*(01), 1950004. https://doi.org/10.1142/S1084946719500043

Hidayah, N., Azhar, Z., Setiany, E., Utami, W., & Tarmidi, D. (2024). Business: Determinants of erm quality and its impact. *Business: Theory and Practice*, *25*(1), 11–23. https://doi.org/10.3846/btp.2024.19302

Hidayah, N., Hapsari, D. W., Adi, K., Saputra, K., Ari, N., Dharmawan, S., & Yadiati, W. (2023). Can institutional good governance and intellectual capital affect university quality? *International Journal of Economics and Management*, *17*(2), 251–261. https://doi.org/10.47836/ijeam.17.2.07

Jayawarsa, A. A. K., Deswari, P. N. N., & Wulandari, I. G. A. (2023). Influence economic growth and percentage of the unemployed population to poverty in Indonesia 2016–2020. *International Journal of Environmental, Sustainability, and Social Science*, *4*(6), 1678–1683. https://doi.org/10.38142/ijesss.v4i6.892

Konovalova, Y. A. (2022). 'New normal': Example of India. In E. G. Popkova & I. V. Andronova (Eds), *Current Problems of the World Economy and International Trade* (Vol. 42, pp. 115–123). Emerald Publishing Limited. https://doi.org/10.1108/S0190-128120220000042012

Li, Y., Wang, M., Liao, G., & Wang, J. (2022). Spatial spillover effect and threshold effect of digital financial inclusion on farmers' income growth – Based on provincial data of China. *Sustainability*, *14*(3), 1838. https://doi.org/10.3390/su14031838

Luwihono, A., Suherman, B., Sembiring, D., Rasyid, S., Kalbuana, N., Saputro, R., Prasetyo, B., Taryana, T., Suprihartini, Y., Asih, P., Mahfud, Z., & Rusdiyanto, R. (2021). Macroeconomic effect on stock price: Evidence from Indonesia. *Accounting*, *7*(5), 1189–1202. https://doi.org/10.5267/j.ac.2021.2.019

Miar, M., & Yunani, A. (2020). The analysis of influence of the government expenditure on poverty in Indonesia. *Jurnal Ekonomi Pembangunan*, *18*(1), 91–102. https://doi.org/10.22219/jep.v18i1.12770

Mokoginta, L. T., Warongan, J. D. L., & Pusung, R. (2024). Analysis of treatment of fixed assets based on government accounting standard statement No. 07 at the Likupang Timur District Office North Minahasa District. *Journal of Governance, Taxation and Auditing*, *2*(4), 229–238. https://doi.org/10.38142/jogta.v2i4.999

OECD. (2021). *OECD economic surveys: Indonesia 2021*. OECD Publishing. https://doi.org/10.1787/fd7e6249-en

Olilingo, F. Z., & Putra, A. H. P. K. (2020). How Indonesia economics works: Correlation analysis of macroeconomics in 2010–2019. *The Journal of Asian Finance, Economics and Business (JAFEB)*, *7*(8), 117–130. https://doi.org/10.13106/jafeb.2020.vol7.no8.117

Pang, J., Jiao, F., & Zhang, Y. (2022). An analysis of the impact of the digital economy on high-quality economic development in China – A study based on the effects of supply and demand. *Sustainability*, *14*(24), 16991. https://doi.org/10.3390/su142416991

Primanto, A., Tanjung, A. M., Situmorang, M., Karim, A. M., Puspitasari, L., Purba, N., & Pramono, R. (2021). Inter-regional cooperation in improving welfare and resolving poverty in Indonesia. *Academy of Strategic Management Journal*, *20*, 1–11.

Priyangga, M. R. Y., & Yasyfi, M. H. (2020). Quo vadis Kebijakan: Analisa Vicious Circle of Poverty Nelayan Tradisional akibat Kebijakan PSBB [Quo vadis policy: Analysis of the vicious circle of poverty of traditional fishermen due to the large-scale social restrictions (PSBB) policy]. *NeoRespublica: Jurnal Ilmu Pemerintahan*, *2*(1), 37–55 (in Indonesian).

Purwono, R., Wardana, W. W., Haryanto, T., & Khoerul Mubin, M. (2021). Poverty dynamics in Indonesia: Empirical evidence from three main approaches. *World Development Perspectives*, *23*, 100346. https://doi.org/10.1016/j.wdp.2021.100346

Qin, C., & Zhang, W. (2022). Green, poverty reduction and spatial spillover: An analysis from 21 provinces of China. *Environment, Development and Sustainability*, *24*(12), 13610–13629. https://doi.org/10.1007/s10668-021-02003-w

Sambodo, M. T., & Novandra, R. (2019). The state of energy poverty in Indonesia and its impact on welfare. *Energy Policy*, *132*, 113–121. https://doi.org/10.1016/j.enpol.2019.05.029

Sameti, M., Esfahani, R. D., & Haghighi, H. K. (2022). Theories of poverty: A comparative analysis. *Arabian Journal of Business and Management Review (Kuwait Chapter)*, *1*(6), 45–56.

Santika, T., Wilson, K. A., Law, E. A., St John, F. A. V., Carlson, K. M., Gibbs, H., Morgans, C. L., Ancrenaz, M., Meijaard, E., & Struebig, M. J. (2021). Impact of palm oil sustainability certification on village well-being and poverty in Indonesia. *Nature Sustainability*, *4*(2), 109–119. https://doi.org/10.1038/s41893-020-00630-1

Saputra, K. A. K. (2021). The effect of sound governance and public finance management on the performance of local governments. *Russian Journal of Agricultural and Socio-Economic Sciences*, *118*(10), 32–43. https://doi.org/10.18551/rjoas.2021-10.04

Saputra, K. A. K., Subroto, B., Rahman, A. F., & Saraswati, E. (2023). Mediation role of environmental management accounting on the effect of green competitive advantage on sustainable performance. *Journal of Sustainability Science and Management*, *18*(2), 103–115. https://doi.org/10.46754/jssm.2023.02.008

Sara, I. M., Udayana Putra, I. B., Kurniawan Saputra, K. A., & Jaya Utama, I. W. K. (2023). Financial literacy, morality, and organizational culture in preventing financial mismanagement: A study on village governments in Indonesia. *Cogent Business & Management*, *10*(1), 1–15. https://doi.org/10.1080/23311975.2023.2166038

Setiawan, S. A. (2020). Does macroeconomic condition matter for stock market? Evidence of Indonesia stock market performance for 21 years. *Jurnal Perencanaan Pembangunan: The Indonesian Journal of Development Planning*, *4*(1), 27–39. https://doi.org/10.36574/jpp.v4i1.105

Soegoto, H., Soegoto, S. W., & Meyer, D. F. (2022). The role of domestic investment, foreign investment and the number of micro small and medium-sized enterprises to reduce poverty in Indonesia. *Journal of Eastern European and Central Asian Research (JEECAR)*, *9*(5), 901–913. https://doi.org/10.15549/jeecar.v9i5.1072

Stoeffler, S. W., & Joseph, R. (2020). Etiology of poverty: A critical evaluation of two major theories. *The Journal of Sociology & Social Welfare*, *47*(1), 73. https://doi.org/10.15453/0191-5096.4269

Sucipto, B., Yusuf, M., & Mulyati, Y. (2022). Performance, macro economic factors, and company characteristics in Indonesia consumer goods company. *Riwayat: Educational Journal of History and Humanities*, *5*(2), 421–428. https://doi.org/10.24815/jr.v5i2.28877

Sugiharti, L., Purwono, R., Esquivias, M. A., & Jayanti, A. D. (2022). Poverty dynamics in Indonesia: The prevalence and causes of chronic poverty. *Journal of Population and Social Studies*, *30*, 423–447. https://doi.org/10.25133/JPSSv302022.025

Suryahadi, A., Al Izzati, R., & Suryadarma, D. (2020). Estimating the impact of COVID-19 on poverty in Indonesia. *Bulletin of Indonesian Economic Studies*, *56*(2), 175–192. https://doi.org/10.1080/00074918.2020.1779390

Tehubijuluw, Z., Yusriadi, Y., Firman, H., & Rianti, M. (2021). Poverty alleviation through entrepreneurship. *Journal of Legal, Ethical and Regulatory Issues*, *24*(1S), 1.

Wang, Q., Guo, J., Li, R., Mikhaylov, A., & Moiseev, N. (2023). Does technical assistance alleviate energy poverty in sub-Saharan African countries? A new perspective on spatial spillover effects of technical assistance. *Energy Strategy Reviews*, *45*, 101047. https://doi.org/10.1016/j.esr.2022.101047

Wang, Z., Ma, D., Zhang, J., Wang, Y., & Sun, D. (2023). Does urbanization have spatial spillover effect on poverty reduction: Empirical evidence from rural China. *Economic Research-Ekonomska Istraživanja*, *36*(1), 2167730. https://doi.org/10.1080/1331677X.2023.2167730

Xu, G., Zhao, T., & Wang, R. (2022). Research on the efficiency measurement and spatial spillover effect of China's regional e-commerce poverty alleviation from the perspective of sustainable development. *Sustainability*, *14*(14), 8456. https://doi.org/10.3390/su14148456

Zhou, Y., & Huang, H. (2023). Geo-environmental and socioeconomic determinants of poverty in China: An empirical analysis based on stratified poverty theory. *Environmental Science and Pollution Research*, *30*(9), 23836–23850. https://doi.org/10.1007/s11356-022-23839-3

Zhu, X., Chen, X., Cai, J., Balezentis, A., Hu, R., & Streimikiene, D. (2021). Rural financial development, spatial spillover, and poverty reduction: Evidence from China. *Economic Research-Ekonomska Istraživanja*, *34*(1), 3421–3439. https://doi.org/10.1080/1331677X.2021.1875859

Changing times, changing drives: Entrepreneurial motivation across generations

Dominik Salat[1], Jarmila Duhacek Sebestova[2]

[1] *Silesian University in Opava, School of Business Administration in Karvina, Department of Business Economics and Management, Czech Republic, ORCID: 0009-0003-0788-9308, salat@opf.slu.cz;*

[2] *Silesian University in Opava, School of Business Administration in Karvina, Department of Business Economics and Management, Czech Republic, ORCID: 0000-0002-7493-0759, sebestova@opf.slu.cz (corresponding author).*

Abstract: *Entrepreneurs from generation to generation play a significant role in the business environment. Understanding this issue, especially by policymakers, is crucial for development. Misunderstanding or lack of interest often leads to inaction or incorrect decisions, creating barriers for current and potential entrepreneurs. Each generation evolves in its own era, and over time, each will fade and be replaced by a new, undefined generation. Policymakers would be interested in understanding these changes and providing targeted support for the needs of different generations, together with removing barriers that arise over time. This paper explores the differences in entrepreneurial motivation. As times change, entrepreneurs must adapt, leading to shifts in behaviour and personality. The study examines motivations for starting businesses across generations, considering variables such as business field, experience, age, sales, and type. Primary research in 446 business entities using the CAWI (computer assisted web interviewing) method tested generational differences in business motivation and participation. The statistical method, Cramer's V, identified interrelationships between variables. Four hypotheses were tested, revealing an influence between necessity motivation and financial results, and a significant tie between generation type and company establishment. These findings could increase interest in this topic, which is expected to revive interest among researchers and policymakers. Simultaneously, the output increases awareness of this issue and brings knowledge to improve the quality of business conditions for all generations.*

Keywords: *Motivation, personality, behaviour, generational differences, entrepreneurship.*

JEL Classification: *L26, M21.*

APA Style Citation: Salat, D., & Duhacek Sebestova, J. (2026). Changing times, changing drives: Entrepreneurial motivation across generations. *E&M Economics and Management*, *29*(2), 45–60. https://doi.org/10.15240/tul/001/2025-5-025

Early Access Publication Date: *December 16, 2025.*

Introduction

For national economies, both entrepreneurs and labour are equally important. Consequently, the best practice for entrepreneurs is to increase investment in innovative activities that generate added value, and for their employees to work more effectively to help the national economy become much more efficient (Ensari, 2017). Entrepreneurship is among the most discussed topics of the last decade. More dynamic and progressive entrepreneurs, according to Zachary and Mishra (2011), could offer a directional path for economic vitality. Currently, a wide variety of forms of entrepreneurial behaviour can be identified, created by the activities of individuals pursuing multiple entrepreneurial opportunities. Douglas et al. (2021) have attempted to solve this challenge by classifying entrepreneurs into types, which are then investigated in models focused on extracting the intensity of entrepreneurial intention for a specific type or a hybrid mixture of types. A growing body of research (Baron, 2006; Rauch & Frese, 2007; Welpe et al., 2012) on social cognitive pathways (e.g., self-efficacy and fear of failure) has contributed to the understanding of the determinants of entrepreneurial behaviour. Despite all these efforts, Murnieks et al. (2014) argue that previous findings on the uniqueness of entrepreneurial psychology are usually inconsistent and produce poor, often inconsistent, or even null outcomes. Despite all the efforts to develop educational schemes and activities to promote entrepreneurship and support ecosystem growth, it is questionable, in the words of Shymko and Khoury (2023), if they will really bring significant transformation and motivate individuals to take up entrepreneurship as a viable career. Understanding these drivers helps the research community to address the complexity of entrepreneurship.

Each individual is unique and changes along with the environment in which they are located. For example, the era of digitalisation has significantly impacted the new entrepreneurial generation and enabled new methods of conducting business that were previously unknown. The aim of this paper is to explore the differences in entrepreneurial motivation across various generations of entrepreneurs. This paper presents a cross-generational view of entrepreneurs and addresses the following research questions:

RQ1: Do entrepreneurial behaviours differ across generations?

RQ2: Do environment and time affect entrepreneurial motivation?

RQ3: Are optimal conditions created for entrepreneurs of all generations?

This research will offer a timely insight into this issue, as changing times necessitate adaptation by many entrepreneurs, potentially leading to a behavioural shift and personality development.

1 Theoretical background

Generational differences have become a central issue in the mainstream press, business literature, scholarly conferences, and workshops, as noted by Oh and Reeves (2011). This is partly due to the lack of standardised terminology used to describe generations, resulting in researchers creating a variety of proprietary terms for each generation. Another reason is the inconsistent range of years defining each generation. Ensari (2017) traced the formation of generational theory in entrepreneurship issues back to the 1950s. Based on this theory, every new generation has divergent working habits. Hence, it is deducted that people born and raised in the same period were also influenced by similar economic, technological, and political trends and therefore exhibit mutual parallels. These premises in generational theory lead to the emergence of generational cohorts (Arkorful et al., 2025; Esteves et al., 2024). For example, Stillman and Stillman (2017) categorised generations and assigned specific general attributes as follows: (i) Traditionalists (born pre-1946), known as “The Silent Generation,” individuals born in this period prioritised traditional ideals of loyalty and patriotism. They were unlikely to challenge authority at work or freely share personal information or emotions; (ii) Baby Boomers (born 1946–1964), also known as the “Me Generation,” they pursued the American dream presented to them. They advocate for idealism, fair rights, and equal opportunities. They often challenge the establishment and are more open to dialogue than previous generations; (iii) Generation X (born 1965–1979), named after the mathematical variable “X” due to being overshadowed by the Baby Boomers during their formative years. This generation has the least definable attributes compared to the adjacent generations. Individuals in this cohort were raised during a period of societal value transformation. For this generation, there is a perceived balance between the quantity and quality of time devoted to parenthood; (iv) Millennials (also known as Generation Y, born 1980–1994). They were raised in a world of economic growth and have never experienced a period without digitalisation. They are often characterised as the first generation of children with structured schedules. Their parents encouraged them to make shared decisions in the home. They are achieving higher educational attainment compared to earlier generations.

Generation Z (born 1995–2012). This generation represents up to 25% of the American population, outnumbering Baby Boomers and Millennials. The majority of this generation are descendants of Generation X, with some being

Tab. 1: Generations and differences of individuals in each generation

Criteria	Silent Generation	Baby Boomers	Generation X	Generation Y	Generation Z
	1920–1945	1946–1964	1965–1979	1980–1994	1995–2012
Job participation	Majority is retired	Some of them are in the declining stage of their career, some of them in the late stage of their career, some of them are retired	Most of them at mid- career and late career stages	Most of them at the the mid-career stage, some at establishment stage of their employment cycle	The majority is student, and a part of it is at exploration stage
Job perspectives	Working for a lifetime employment and safety	Living to work	Working to live	Balance of work and life	Flexible lives Flexible works
Working life	The job is important for the needs	Workaholic	Work-life balance is important	Putting some fun in the work	Not known clearly yet
Authority	Loyal	Loyal	Questions the authority	Rejects the authority	Find the authority unnecessary
Working hours	Works long hours	Works long hours	Wants to work with flexible working hours	Wants flexible working hours	Not known clearly yet
The reward expectation from the workplace	Financial security	Lifetime employment	Intangible subjects increase the job satisfaction	Job changing habits are high	Advancement is important
Technology	Weak in technology	Weak in technology	The use of technology and the internet is good	Born in technology and the internet	Born in technology and the internet

Source: Ensari (2017)

children of Millennials. They are in harmony with the physical and digital environment. They were born and raised after 9/11 and through the recession. Increasing global terrorism and insecurity have become a normal aspect of daily life. Csobanka (2016) also describes the Alpha generation (beginning in 2013) as materialistic. The differences between these generations are illustrated in Tab. 1.

One factor, a relatively small variable, that usually differentiates generations is the year of birth, according to Bejtkovsky (2016). More importantly, he suggests the formulation of generations is better defined by the historical events they have lived through rather than by chronological periods. With increasing age, individuals change gradually, experiencing a change in social role or social status. They lose physical strength, which contributes to a decline in economic strength. Gradually, they begin to consider retirement and to realise the length of their life. Seki et al. (2022) suggest that these age-related changes in older people shape their unique personality traits. In this context, Baltes et al. (1997) identified three evolutionary pathways that highlight two critical points in the life cycle where individuals are susceptible to environmental imprinting.

Firstly, the evolutionary pathway suggests that biological plasticity and the ability to respond to important environmental signals decrease with age. This assertion is based on existing knowledge of DNA-conditioned changes in age-specific expression, which demonstrate a reduction in the ability of natural selection to operate with increasing age (Martin et al., 1996). According to these findings, individuals are more prone to environmental influences in infancy. Biological adaptability reaches its maximum as a result of differences in the genes in response to environmental conditions. These changes are phenotypical because they transform the way a gene is normally expressed and do not result in a corresponding alteration of the genotype alone.

Secondly, according to Baltes et al. (1999), the cultural pathway suggests that increasing age influences the need for resources (psychological, social, material, and knowledge, such as technology and economics), as the biological ability to adapt to the characteristics of the environment declines.

Finally, the third aspect, also a cultural pathway, notes the diminishing effectiveness of culture with increasing seniority. Taken together, these cultural pathways reflect an expanding requirement for culture because of biological decline, and a drop in culture effectiveness at later life phases as biological decline proceeds. The cultural aspect demonstrates a likely tendency for environmental imprinting during young adulthood (between 18 and 25 years of age), as cultural demands, the necessity for culture, and the effectiveness of culture are in a communal "sweet spot" of transition. During young adulthood (which is regarded as a formative time in life), individuals become proactive players in the social environment and enter into interaction with a broader set of different sources of knowledge that form adopted standards, norms, values, and attitudes (Arnett, 2000). Tooby and Cosmides (1990) suggested that the combination of learning outcomes, evolutionary plasticity, and neurodevelopmental adjustment to the environment via reactions to signals from the habitat or situational appraisals result in multiple transitory and enduring differences in individually determined behaviours, including variations in personality. Characteristic of Homo sapiens, human naturalness indicates unique neurologic, cognitive, and psychological adjustments and propensities, as similarly argued by Maestripieri and Boutwell (2022).

Social cognitive theory (Bandura, 1986; Lent & Brown, 1996) posits that human behaviour is affected by personality factors, behavioural factors, and environmental factors, which mutually impact in a tripartite relationship. This is considered by Baron (2004) to be the nouveau area in the entrepreneurship literature. Entrepreneurial cognition (Mitchell et al., 2002, p. 97) can be defined as "*knowledge structures that people use to make assessments, judgments, or decisions involving opportunity evaluation, venture creation, and growth.*" This uniqueness is particularly important because all individuals are fundamentally unique and cannot be copied. The difference is noticeable in their performance, as every individual aims to maximise their advantages and outperform the rest. Consequently, they differ from each other. Depending on what they excel at, they choose their next path to follow. For this reason, social psychologists are intensely investigating the predisposing power of personality traits. Thus, according to Ness et al. (2020), they seek to specify information regarding the impact of personality on the manner in which individuals experience and negotiate with the outside elements of their personal circumstances and workplace scenarios.

The necessary knowledge relates to the identification of possibilities for private, career, or trade activities, as well as the broader context of the environment in which people live or work, and the knowledge of the functioning of the economy and the opportunities and challenges confronting employers or organisations. Therefore, it is also up to the public to be aware of the ethical status of companies and how companies can act to benefit society through fair business or social entrepreneurship. Proactively managing a project includes the skills of planning, organising, managing, leading people and delegating, analysing, communicating, evaluating, and reporting. Communication and negotiation skills are also important, along with the ability to work individually and in a team. It is also essential to assess and identify strengths and weaknesses and demonstrate tolerance of higher risk if necessary. The entrepreneurial mindset is defined by being proactive, autonomous, and innovative in both private and public situations as well as in the workplace. It also involves motivation and commitment to goals, whether individual or shared, including professional goals (Kurek & Rachwa, 2011).

Ness et al. (2020), among all the ambiguous findings, assume that business founders possess personality traits specific to them. It is quite likely that the inconsistency of the research results can be attributed to the different manifestations of the entrepreneurial samples investigated. In other words, different personality profiles are linked to various personality characteristics for every type of entrepreneurial involvement. In their most recent paper, Develi et al. (2011, p. 118) argue that "*The most important characteristic of the entrepreneurial personality is the tendency to take risks or the predisposition to take risks.*" Related to this, the percentage of seniors who start a new venture is approximately less than double the rate of young people who start a new enterprise (Kautonen et al., 2008). Conversely, research conducted by the Global Entrepreneurship

Monitor (GEM, 2025) reports that young entrepreneurs survive for longer than the first three and a half years, while adults over the age of 34 are 1.7 times more likely than young people to run mature enterprises. Young people form up to 73% of unincorporated entities (GEM, 2025) in comparison to other generational entrepreneurs. This can stem from the higher priority given to income, security factors, and flexibility (Waschkewitsch & Kuhnt, 2025) for Generation Z compared to other generations. This generation also has a strong need for feedback (Dwivedula, 2025; Zieger, 2024). Compared to other cohorts, Generation Z shows less interest in employment stability; they frequently change their jobs, seeking flexibility and an escape from the daily routine. Generation Z identifies self-employment as a method of professional empowerment, primarily because they believe it offers higher earnings and a better perception of autonomy (Dolot, 2018). Conversely, Generation X is more likely to choose to work in larger companies or start their own businesses (Campbell et al., 2017). Despite these differences in approach, the underlying goals remain consistent across All generations. Across generations, a common feature is the desire to improve their financial condition (Fuchs et al., 2024; Perez-Encinas et al., 2021). Furthermore, from Generation Y onwards, personal development is considered one of the most important motivation criteria (Fuchs et al., 2024).

According to Lévesque and Minniti (2006), older adults are less willing to participate in entrepreneurship due to health conditions, time scheduling preferences, or other additional motives. The study by Kautonen et al. (2014) reported that the age-entrepreneurship propensity-correlated relationship depends largely on the heterogeneous preferences of individuals toward their professional life. The relationship between age and entrepreneurial propensity is described by an inverted U-shaped curve, which peaks at age 48 for individuals who prefer to establish their own ventures and invest time and capital in them. This is consistent with the traditional view of a declining propensity for entrepreneurship among older adults. Conversely, for individuals pursuing self-employment who do not intend to hire new staff, the propensity to be entrepreneurial increases linearly with age. Reluctant entrepreneurs (those driven into entrepreneurship by necessity) exhibit a marginal effect of age on their entrepreneurial activities. Harms et al. (2014) additionally introduce the category of "silver-age" entrepreneurs.

2 Research methodology

The theoretical review informed the selection of research topics for quantitative research, such as generational motivation to start up (Kautonen et al., 2014). The descriptive variables included the field of business, the length of business experience, the age of the respondent, sales, and the type of business entity. The independent variables were the motivation for entrepreneurship and the life cycle of the business. The key factor in the allocation to generations was the equal time interval in which these individuals were born. The Stillman and Stillman (2017) classification was used to divide the respondents into different generational cohorts. According to this classification, based on the age of the person responding as well as the year of birth, they were classified into different generational cohorts as characterised by these authors.

Situation within the Czech Republic. According to the Trading Economics (2023), self-employed persons, total in the Czech Republic was reported at 16.49% of total employment in 2023. According to the Czech Statistical Office (2024), there were 1,291,474 in total active businesses in the Czech Republic. The composition of the generation structure was based on a detailed analysis of the Czech Statistical Office from 2017. Tab. 2 illustrated potential changes in 2025.

Generation Z and Younger Millennials represent a smaller, but potentially growing segment of self-employed individuals. They are less likely to be in traditional "trade licence" self-employment and more involved in flexible, project-based work. Older Millennials are highly active in entrepreneurship, especially in their mid to late 30s and early 40s. Generation X forms a substantial core of self-employed individuals in the Czech Republic, often in their peak career years. Finally, Baby Boomers show a significant and growing presence in self-employment, particularly as they reach traditional retirement ages but continue to work. Understanding cohort bias is therefore crucial for this classification. Dvouletý et al. (2024) note that current economic conditions, post-retirement life, and accumulated skills and experiences

Tab. 2: Generational changes within active business entities

Age group (2017)	Approximate age in 2025	Generation name	Share of self-employed individuals within their age groups in the total working population in 2017 (%)
Up to 29 years	16–34 years old	Primarily Generation Z (those born 1997–2009, 16–28 years old in 2025) and youngest Millenials (1991 to 1996)	9.2
30 to 39 years	38–47 years old	Millenials	31.3
40 to 49 years	48–57 years old	Generation X	37.6
50 or more years	58–68+ years old	Baby Boomers	65.0

Source: Czech Statistical Office (2018)

from a particular historical trajectory are important for entrepreneurial generations. This highlights a limitation of this study.

2.1 Conceptual research framework and hypotheses development

Subsequently, each respondent was classified as one of the types of entrepreneurs. The data obtained were then used to cluster into homogeneous groups. According to Douglas et al. (2021), empirical models typically combine the motivation to be an entrepreneur with the motivation to be a certain type of entrepreneur. Therefore, most previous studies focused on a particular type of entrepreneur and tried to explain the determinants of this specific type of behaviour. For this reason, it was not necessary to perform a cluster analysis, but it was possible to use the categories already created. In this way, the research was able to build on research already conducted on this aspect. Therefore, the breakdown of entrepreneurs as described by Zhang and Acs (2018) in their publication was selected.

Based on this, the first type represented novices vs. non-novices, where different propensities to entrepreneurship might help explain the relationship between age and entrepreneurship. By definition, younger people were more likely to be budding entrepreneurs with limited prior entrepreneurial experience. For many older workers, after decades of working for others, they might decide to become budding entrepreneurs. The accumulation of work experience and wealth at older ages increased physical, human, and social capital and thus increased entrepreneurial opportunity. Therefore, we expected more opportunities than necessity for entrepreneurs as people aged. Therefore, other types were defined as opportunity or necessity entrepreneurs. Due to the increased responsibility to care for children and family around age 30, many workers, particularly women, changed to work part-time; we thus expected a decline in the full-time (versus part-time) entrepreneur rate around the age of 30. At later life, with needs to care for a spouse, their own health, or a more relaxed lifestyle, phased retirement arose and we expected a continued declining full-time (versus part-time) entrepreneur rate. While sole proprietorship increased with age, the role increased first (until late 40s in Europe) and then decreased with age. Since incorporated entrepreneurs typically operated larger businesses with paid employees than unincorporated entrepreneurs, sole proprietors were more likely to be unincorporated entrepreneurs; owner-managers were more likely to be incorporated entrepreneurs. Therefore, we expected that the propensity to be an incorporated (versus unincorporated) entrepreneur increased first and then decreased. Through this distribution, it was possible to identify the dominant generation in each type of entrepreneurs. Although most of Generation Z was expected in the case of the novice type, further results were sought. This allowed for a different perspective on each generation and their enterprise. Due to the lack of representation of the Silent Generation, the sample of this generation was not further researched. A conceptual model with individual variables was created (Fig. 1), where in this case the dependent variable was Generations (x) and the independent variables were Stages of the enterprise life cycle, Pressure from the surrounding area and Improvement of the financial situation, Sales earned by the company (CZK/year) and Establishment of a company (y). In addition, there was a third variable in the model. This was represented by the Duration of the entrepreneurial experience, another Motivation to do business, Stages of the enterprise life cycle and Pressure from the surrounding area and Improvement of the financial situation (z).

Since the entrepreneurs were divided into different types only on the basis of deduction, it was necessary to verify this statement through the hypothesis:

H1: The relationship between entrepreneurial type (novices vs. non-novices) and Stages of the enterprise life cycle is moderated by the Duration of entrepreneurship experience across generational cohorts.

H2: Choosing opportunity vs. necessity across Generations is negatively associated with Pressure from the surrounding area and Improvement of the financial situation, depending on another Motivation to do business.

H3: Full-time vs. part-time operation across Generations is positively influenced more towards full-time due to Sales earned by the company (CZK/year) in relation to Stages of the enterprise life cycle.

H4: The association between the choice of incorporated vs. unincorporated business structure and the Establishment of a company is moderated by the combined influence of Pressure from the surrounding area and Improvement of the financial situation across generational cohorts.

The latter represented the control variable and examined the cumulative effect (Rubin & Babbie, 1997) of the dependent and independent variables in each row of the model. All variables were treated as lower-order variables, i.e., as nominal data. The statistical program SPSS was used to confirm or reject hypotheses, for which Cramer's V was tested. The limitations of the study lay in the fact that it was not possible to accurately determine the distribution by generation in the Czech business environment. The effect of generations was expected to be seen in their view of the financial side of business.

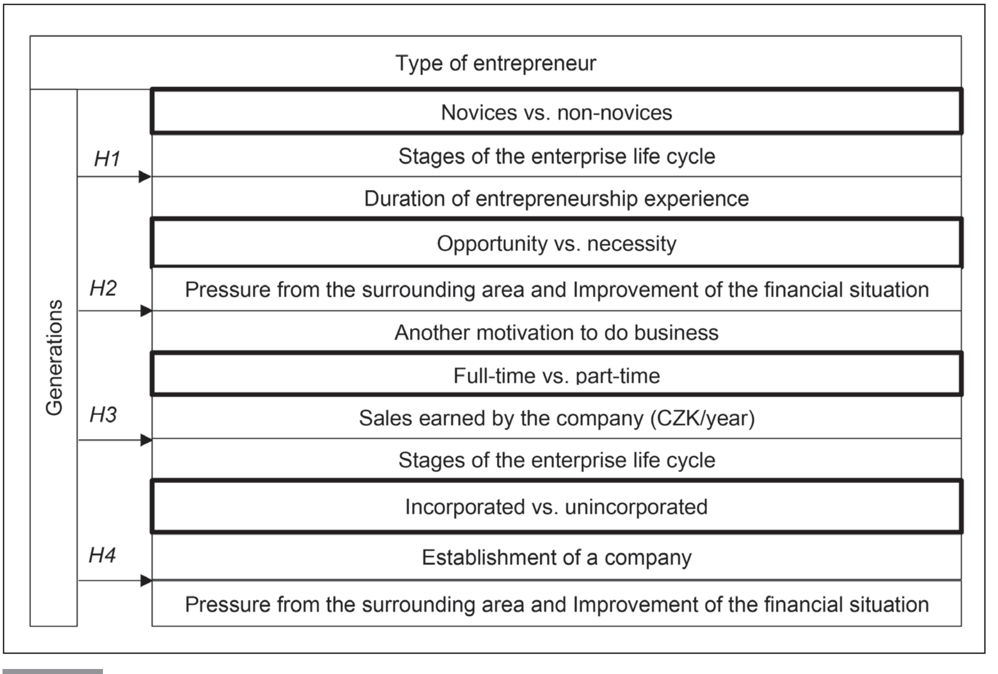

Fig. 1: Conceptual model of entrepreneurial type

Source: Zhang and Acs (2018), own modification

2.2 Data sample description

In this primary research, a questionnaire was used for data collection. This method was considered one of the most effective for collecting a large volume of data from many persons in a relatively short period of time, as the CAWI (computer-assisted web interviewing) method was employed. These persons were enterprise entities operating on the territory of the Czech Republic. In total, 446 subjects finally participated in this study, and primary data were collected between October and December of the year 2023. The questionnaire was structured in two parts: the first addressed questions on attitudes in entrepreneurship, and the second section focused on the respondent's profile. The structure of the subjects was randomly selected from the business database Amadeus. The obtained data sample followed the original structure of the business population within the Czech Republic according to current statistics (MPO, 2024). More information and characteristics of the respondents are presented

Tab. 3: Respondent's description – Part 1

Variable	Groups	Sample (N = 446)	Sample (%)
Generations	Silent Generation	0	0.0
	Baby Boomers	62	13.9
	Generation X	198	44.4
	Generation Y	123	27.6
	Generation Z	63	14.1
Establishment of a company	Unincorporated entity (sole proprietors)	257	57.6
	Incorporate entity	189	42.4
Duration of entrepreneurship experience	Less than 3 years	87	19.5
	From 3 to 10 years	130	29.2
	From 11 years to 20 years	87	19.5
	Over 20 years	142	31.8
Stages of the enterprise life cycle	Start-up company	52	11.7
	Company in growth	276	61.9
	Company in stagnation	104	23.3
	Company in decline	9	2.0
	Company is going into liquidation	5	1.1
Field of enterprise	Manufacturing	70	15.7
	Trade	85	19.1
	Services	279	62.6
	Agriculture	5	1.0
	Other	7	1.6
Sales earned by the company (CZK/year)	Up to 49 million	387	86.8
	From 50 million to 249 million	41	9.2
	From 250 million to 1.24 billion	9	2.0
	Over 1.24 billion	9	2.0

Tab. 3: Respondent's description – Part 2

Variable	Groups	Sample (*N* = 446)	Sample (%)
Motivation to do business	Pressure from the surrounding area (NM)	45	10.0
	Try something different (OM)	98	22.0
	Improvement of the financial situation (NM)	99	22.2
	Regional or national development (OM)	20	4.5
	The satisfaction of achieving something great (OM)	49	11.0
	Leave a legacy for the future generation (NM)	19	4.3
	Personality development (OM)	92	20.6
	Talent development (OM)	24	5.4

Note: NM – necessity motivation; OM – opportunity motivation; exchange rate 1 Euro = 24.19 CZK.

Source: own (based on primary data)

in Tab. 3. The size of enterprises was classified according to the EU definition of small and medium enterprises, where micro enterprises employed fewer than 10 persons (annual turnover or annual balance sheet total not exceeding EUR 2 million); small enterprises employed fewer than 50 persons (annual turnover or annual balance sheet total not exceeding EUR 10 million); and medium-sized enterprises employed fewer than 250 persons (annual turnover not exceeding EUR 50 million or an annual balance sheet total not exceeding EUR 43 million) (European Commission, 2020).

The classification of a full-time or part-time entrepreneur was based on three factors: generation, Sales earned by the company (CZK/year; exchange rate 1 Euro = 24.19 CZK), and Stages of the enterprise life cycle. In the case of Generations, entrepreneurs were classified as full-time if their sales were CZK 49 million or more and their enterprise was not in stagnation, decline, or liquidation. Otherwise, if the entrepreneur had sales below CZK 49 million and was not operating a start-up or growth company, such an entity was classified as a part-time entrepreneur. Regarding the opportunity or necessity distinction, the entrepreneur was classified as a necessity entrepreneur when their main motive for doing business was Pressure from the surrounding area and Improvement of the financial situation. Conversely, when the motive was something else, the entrepreneur was classified as an opportunity entrepreneur. However, this classification was tested statistically. The testing findings are presented in the results section.

3 Results and discussion

This section presents, first, the results of the hypothesis testing, followed by a discussion of these findings in the context of existing literature.

3.1 Results

The results for individual hypotheses, *H1* through *H4*, are now presented. It begins by detailing the specific statistical test employed for the evaluation of each hypothesis.

Subsequently, each finding is thoroughly assessed and interpreted with direct reference to the evidence gathered through the survey questionnaire, ensuring a robust link between our methodology and the empirical results.

Enterprise life cycle and entrepreneurship experience (*H1*). Initially, the association between Generations and the Stages of the enterprise life cycle was tested, and a strong, statistically significant association was found (Cramer's V = 0.713, Sig. = 0.001). The distribution across life cycle stages revealed clear generational patterns: companies associated with Generation Z were predominantly (83.9%) start-up ventures, whereas all representatives (100%) of Generation X were companies in the growth stage. For Generation Y, the companies were split between growth (55.3%) and stagnation (44.7%), and among Baby Boomers, the majority (77.8%) were in stagnation, with 14.3% in decline and 7.9% moving into liquidation. Furthermore, the association between generations and the Duration of entrepreneurship experience also demonstrated a strong, statistically significant association (Cramer's V = 0.706, Sig. = 0.001). Data indicated major differences in experience, with 100% of Generation Z entrepreneurs having less than three years of experience, compared to 100% of Baby Boomers who had more than 20 years of experience. Generation X was comprised of 65.7% of entrepreneurs with 3 to 10 years of experience, 21.7% with 11 to 20 years, and 12.6% with less than 3 years, while for Generation Y, 64.2% had more than 20 years of experience and 35.8% had 11 to 20 years. When controlling for the relationship with the third variable (Duration of entrepreneurship experience), Cramer's V remained high (Cramer's V = 0.653, Sig. = 0.001), and no additive effect, spurious association, or hidden association was found. Therefore, hypothesis *H1* was confirmed.

Generation type and start-up motive: Pressure from the surrounding (*H2*). When examining the association between Generations and the combined variable of Pressure from the surrounding area and Improvement of the financial situation, a strong, statistically significant association was found (Cramer's V = 0.692, Sig. = 0.001). This primary motivation (necessity) was identified almost exclusively among Baby Boomers (100%). Within this cohort, the necessity was split between Pressure from the surrounding area (72.6%) and the Improvement of the financial situation (27.4%). The Improvement of the financial situation was also observed in Generation X (41.4%), while for the rest of that generation (58.6%), another motivation for business was dominant. Conversely, for both Generation Y and Z, the primary motivation was another Motivation to do business (100% in both cases). Consequently, the direct association between Generations and another Motivation to do business also yielded a strong, statistically significant relationship (Cramer's V = 0.701, Sig. = 0.001). The most common secondary motivations for Generation X were leaving a legacy for the future generation (9.6%), personality development (37.4%), and the satisfaction of achieving something great (11.6%). For Generation Z, the motivation was solely "try something different" (100%), while Generation Y cited talent development (19.5%), personality development (14.6%), and 65.9% had different, various motivations for their business. Controlling for another Motivation to do business saw Cramer's V rise to 0.724 (Sig. = 0.001), thus increasing the association with the opportunity vs. necessity entrepreneurial type. Therefore, hypothesis *H2* was confirmed.

Generation type, sales and life cycle (*H3*). The association between Generations and Sales earned by the company (CZK/year) was examined, and a statistically significant relationship was found (Cramer's V = 0.556, Sig. = 0.001). The distribution of sales showed a distinct pattern for Baby Boomers, where 6.3% had sales up to CZK 49 million, 65.1% ranged from CZK 50 million to CZK 249 million, 14.3% ranged from CZK 250 million to CZK 1.24 billion, and 6.3% exceeded CZK 1.24 billion. In contrast, all other generations (Generations X, Y, and Z) collectively reached 100% for the lowest sales category (up to CZK 49 million). The results concerning the association between Generations and the Stages of the enterprise life cycle (Cramer's V = 0.713, Sig. = 0.001) were consistent with previously presented findings and were thus used as a control. The association between the independent variables and the control variable (Stages of the enterprise life cycle) reached Cramer's V = 0.685 (Sig. = 0.001), and consequently, an additive effect was found. Therefore, hypothesis *H3* was confirmed.

Generation type and start-up motive: Pressure from the surrounding financial situation improvement (*H4*). A very strong, statistically significant relationship was found between Generations and the Establishment of a company (incorporated vs. unincorporated status) (Cramer's V = 0.986, Sig. = 0.001). This association indicated a clear generational dichotomy in legal structure: Baby Boomers were exclusively found as an incorporated entity (100%), whereas Generation Z exclusively chose an unincorporated entity (100%). Generation X predominantly operated as an unincorporated entity (98.5%) compared to a negligible proportion of incorporated entities (1.6%). Conversely, Generation Y was entirely represented by the incorporated entity type (100%). The association between Generations and Pressure from the surrounding area and Improvement of the financial situation was consistent with previous findings (Cramer's V = 0.692, Sig. = 0.001) and served as the control variable. The association with this control variable yielded a statistically significant result (Cramer's V = 0.592, Sig. = 0.001), with 100% of the control variable respondents operating as an unincorporated entity. Therefore, hypothesis *H4* was confirmed.

A summary of all results is presented in Tab. 4, where they are connected with each hypothesis.

Tab. 4: Summary of tests related to hypotheses

Association between variables	Cramer's V	Sig.	Related to
Generations and Duration of entrepreneurship experience	0.706	0.001	*H1*
Stages of the enterprise life cycle and Duration of entrepreneurship experience	0.653	0.001	
Generations and Pressure from the surrounding area and Improvement of the financial situation	0.692	0.001	*H2*
Generations and another Motivation to do business	0.701	0.001	
Pressure from the surrounding area and Improvement of the financial situation and another Motivation to do business	0.724	0.001	
Generations and Stages of the enterprise life cycle	0.713	0.001	*H3*
Generations and Sales earned by the company (CZK/year)	0.556	0.001	
Sales earned by the company (CZK/year) and Stages of the enterprise life cycle	0.685	0.001	
Generations and Establishment of a company	0.986	0.001	*H4*
Establishment of a company and Pressure from the surrounding area and Improvement of the financial situation	0.592	0.001	

Note: *p*-value = 0.05.

Source: own based on primary data

3.2 Discussion

Unlike the broad study conducted by Ensari (2017), this research focuses narrowly on "motivation" as a single factor that evolves over time, driven by changes in an individual's personality development and subsequent entrepreneurial cognition. The entrepreneur's age and the operational period of the business are integral to this analysis, as the influence of these variables cannot be fully separated from that of the generational cohorts (Schröder, 2023). Crucially, Generation Z entrepreneurs are navigating a fundamentally different technological era (Dolot, 2018), enabling them to more effectively leverage the potential of current technologies (Dolot, 2018; Ensari, 2017). The findings of the actual research replicate the results of the Dvouletý study (2019), confirming the dominance of entrepreneurs within the Generation X and Generation Y age cohorts in the Czech Republic. These findings also align with earlier international studies: Generation Z's preference for unincorporated entities (Dolot, 2018; GEM, 2025) grants them greater operational freedom because their primary goal is not leadership (Deloitte, 2025), contrasting sharply with other generations. Conversely, the propensity of Baby Boomers to pursue entrepreneurship is often motivated by a desire for financial stability or a response to perceived age-related discrimination or reduced health (Perez-Encinas et al., 2021). Their strong preference for incorporated entities allows them to fully leverage their leadership skills and accumulated experience, enabling them to serve effectively as mentors and contributors to society (Arkorful et al., 2025).

A summary of the type of entrepreneurs in relation to generation is presented in Fig. 2.

Type of entrepreneurs		Generations					Type of entrepreneurs
	Novices	Z	X	Y	Baby Boomers	Non-novices	
	Opportunity	Z	Y	X	Baby Boomers	Necessity	
	Full-time	X	Z	Y	Baby Boomers	Part-time	
	Incorporated	Baby Boomers	Y	X	Z	Unincorporated	

Fig. 2: Type of entrepreneurs in relation to generation (Gen)

Source: own based on primary data

Fig. 2 represents a general classification of individual entrepreneurs within the sampled population (N = 446), extending the presented studies (Douglas et al., 2021, Zhang & Acs, 2018). It can be posited that the primary differentiating factors between generations are the motivation for entrepreneurship and the temporal commitment each generation is willing to dedicate to their venture. For the Baby Boomer generation, the driving forces are predominantly necessity and the fear of financial insecurity, resulting in a tendency towards part-time entrepreneurship. Generation X, conversely, represents a typical succession cohort that often inherits established business structures. In sharp contrast, Generations Y and Z exhibit a more progressive and opportunity-driven orientation, characterised by a greater inclination toward operational independence and autonomy.

The major limitation of this study, as acknowledged in the methodology, is the difficulty in precisely determining the generational distribution in the Czech business environment, which may affect the generalizability of the specific percentage data. However, the generalizability of this classification is inherently limited by the sample's composition and size. To achieve a more robust and accurate representation of generational distribution, future research would require a substantially larger number of subjects and a greater proportion of the overall business population. Future research should focus on longitudinal studies to track whether Generation Z's initial unincorporated structures and start-up status evolve into the incorporated, growth-oriented firms seen in Generation Y and Generation X, providing clarity on the long-term impact of digitalisation and career priorities on entrepreneurial development. Despite these methodological limitations, the findings hold significant practical implications for the Czech context, particularly in informing the potential focus of business support, the strategic work of supporting infrastructure, and the tailoring of entrepreneurship education across different age cohorts.

Conclusions

Based on the primary research conducted, this study validated that generational affiliation is a defining factor in entrepreneurial involvement and motivation within the Czech Republic. The results indicate a distinct separation: necessity-driven entrepreneurship is concentrated among Baby Boomers, while opportunity-driven ventures characterise Generations Z and Y. This not only extends existing literature but also precisely delineates the current state of Czech entrepreneurs. Additional indicators, such as duration of entrepreneurial experience and turnover generated, facilitated the precise integration of generational cohorts into specific entrepreneurial typologies.

It is evident that Baby Boomers, while highly experienced, are increasingly shifting towards part-time entrepreneurial engagement, reflecting the influence of age-related variables, including declining physical or cognitive health. This necessitates a nuanced policy response. Given that every entrepreneur possesses unique personality traits and entrepreneurial cognition, it is imperative for policymakers to avoid erecting generational barriers and to holistically adapt support conditions. Specifically, the government must prevent the disadvantage of certain cohorts (e.g., Traditionalists) simply because their skills (e.g., technological proficiency) are less advanced than those of younger generations. Government strategy should actively contribute to entrepreneurship development by intentionally working with diverse generational needs. While a broad range of skills development currently exists for seniors, specific resources must be targeted towards “entrepreneurial seniors.” Identifying the precise needs and aspirations of this demographic through dedicated research should be the preliminary step in establishing a comprehensive support process. This strategic approach would enhance the quality of support provided and serve as an additional incentive for older entrepreneurs to maintain activity, adapting effectively as times change. Based on the findings, several customised policy recommendations are warranted:

(i) Tailored support: support programmes should be customised across a diverse portfolio of public and private support mechanisms, respecting the specific needs of each generation. This aligns with the concepts proposed by Douglas et al. (2021), advocating for start-up resources for Generation Z, growth support for Generation X, and revitalisation strategies for Baby Boomers;

(ii) Knowledge transfer ecosystem: to bolster the entrepreneurial ecosystem, experience-based mentorship and flexible education programmes should be implemented. This would effectively extend the experience of Baby Boomers to mentor younger entrepreneurs, fostering crucial knowledge transfer and potentially reducing start-up failure rates, thereby extending the idea proposed by Ness et al. (2020);

(iii) Motivation-driven policies: finally, national and regional policies must be motivation-driven. This entails designing support that respects the unique motivational profiles of each generation—starting with innovation grants for Generation Z, aiming to test new ventures and culminating in security financial support for Baby Boomers.

The findings from this study, particularly the stark differences in entrepreneurial motivation, firm structure, and life-cycle stage, hold immediate implications for Czech national and regional policy, especially within the context of the National Research and Innovation Strategy for Smart Specialisation (RIS3, Ministry of Education, Youth and Sports of the Czech Republic, 2021). Current public support mechanisms often struggle with a one-size-fits-all approach, failing to account for the reality that the Baby Boomer cohort predominantly seeks financial security and revitalisation, while Generation Z requires support focused on rapid innovation, market testing, and digital scaling. The related operational programme, the Just Transition Operational Programme (OP ST) (Ministry for Regional Development, 2021) indicates that policy must become motivation-driven. This means reallocating resources away from uniform subsidies toward customized support packages (Douglas et al., 2021). For instance, providing security-focused grants or micro-loans to sustain older, established businesses (Baby Boomers) and offering low-barrier innovation vouchers or proof-of-concept grants to encourage younger entrepreneurs (Generation Z) to try new, technologically intensive ventures (Ness et al., 2020). The Czech policy framework must recognize that since generational effects cannot be separated from cohort effects, these segmented needs are not transient but are structural features of the entrepreneurial landscape, requiring specific attention within regional action plans derived from the Action Plan for the National RIS3 Strategy (Ministry of Industry and Trade of the Czech Republic, 2021).

Acknowledgements: *This paper was supported by the Ministry of Education, Youth and Sports Czech Republic within the Institutional Support for Long-term Development of a Research Organization in 2023 [IP/05/2023] and Student Research Project "Psychological Capital as a Tool for Developing Entrepreneurial Personality" [SGS/11/2025].*

References

Arkorful, H., Hilton, S. K., & Awaah, F. (2025). Generational cohorts and their predisposition toward entrepreneurship in an emerging economy. *Journal of Economic and Administrative Sciences, 41*(1), 245–262. https://doi.org/10.1108/JEAS-09-2021-0194

Bandura, A. (1986). The explanatory and predictive scope of self-efficacy theory. *Journal of Social and Clinical Psychology*, *4*(3), 359–373. https://doi.org/10.1521/jscp.1986.4.3.359

Baron, R. A. (2004). The cognitive perspective: A valuable tool for answering entrepreneurship's basic "why" questions. *Journal of Business Venturing*, *19*(2), 221–239. https://doi.org/10.1016/s0883-9026(03)00008-9

Baron, R. A. (2006). Opportunity recognition as pattern recognition: How entrepreneurs connect the dots to identify new business opportunities. *Academy of Management Perspectives*, *20*(1), 104–119. https://doi.org/10.5465/AMP.2006.19873412

Bejtkovsky, J. (2016). The employees of Baby Boomers Generation, Generation X, Generation Y and Generation Z in selected Czech corporations as conceivers of development and competitiveness in their corporation. *Journal of Competitiveness*, *8*(4), 105–123. https://doi.org/10.7441/joc.2016.04.07

Campbell, S. M., Twenge, J. M., & Campbell, W. K. (2017). Fuzzy but useful constructs: Making sense of the differences between generations. *Work, Aging and Retirement*, *3*(2), 130–139. https://doi.org/10.1093/workar/wax001

Csobanka, Z. E. (2016). The Z Generation. *Acta Educationis Generalis*, *6*(2), 63–76. https://doi.org/10.1515/atd-2016-0012

Czech Statistical Office. (2018). *Share of the self-employed in total employment in the Czech Republic is one of the highest in the EU28*. https://csu.gov.cz

Czech Statistical Office. (2024). *Business demography statistics*. https://csu.gov.cz/business-demography-statistics

Deloitte. (2025). *Deloitte global Gen Z and Millennial survey 2025*. https://www.deloitte.com/global/en/issues/work/genz-millennial-survey.html

Develi, E. I., Erdil, B., & Sevimli, Y. (2011). Entrepreneurship and importance of personality on entrepreneurship: A research on trainees of entrepreneurship education program. *International Journal of Business and Management Studies*, *3*(1), 115–124.

Dolot, A. (2018). The characteristics of Generation Z. *E-Mentor*, *74*(74), 44–50. https://doi.org/10.15219/em74.1351

Douglas, E. J., Shepherd, D. A., & Venugopal, V. (2021). A multi-motivational general model of entrepreneurial intention. *Journal of Business Venturing*, *36*(4), 106107. https://doi.org/10.1016/j.jbusvent.2021.106107

Dvouletý, O. (2019). Development of entrepreneurial activity in the Czech Republic over the years 2005–2017. *Journal of Open Innovation: Technology, Market, and Complexity*, *5*(3), 38. https://doi.org/10.3390/joitmc5030038

Dvouletý, O., Svobodová, I., Bočková, N., & Šebestová, J. (2024). Becoming a first-time entrepreneur in your 40s and older: Lessons from survival analysis. *World Journal of Entrepreneurship Management and Sustainable Development*, *20*(2). https://doi.org/10.47556/j.wjemsd.20.2.2024.4

Dwivedula, R. (2025). Why Is Generation Z motivated at work? A qualitative exploration. *Global Business and Organizational Excellence*, *44*(3), 38–49. https://doi.org/10.1002/joe.22273

Ensari, M. S. (2017). A study on the differences of entrepreneurship potential among generations. *Research Journal of Business and Management*, *4*(1), 52–62. https://doi.org/10.17261/Pressacademia.2017.370

Esteves, J., de Haro Rodríguez, G., Ballestar, M. T., & Sainz, J. (2024). Gender and generational cohort impact on entrepreneurs' emotional intelligence and transformational leadership. *International Entrepreneurship and Management Journal*, *20*(2), 1295–1322. https://doi.org/10.1007/s11365-024-00955-9

European Commission. (2020). *User guide to the SME definition*. Publications Office of the European Union.

Fuchs, O., Lorenz, E., & Fuchs, L. (2024). Generational differences in attitudes towards work and career: A systematic literature review on the preferences of Generations X, Y and Z. *International Journal of Innovative Research and Advanced Studies*, *11*(7), 54–71.

GEM. (2025). Youth are more entrepreneurial than adults: GEM report on youth entrepreneurship. *GEM Global Entrepreneurship Monitor*. https://www.gemconsortium.org

Harms, R., Luck, F., Kraus, S., & Walsh, S. (2014). On the motivational drivers of gray entrepreneurship: An exploratory study. *Technological Forecasting and Social Change*, *89*, 358–365. https://doi.org/10.1016/j.techfore.2014.08.001

Ministry of Industry and Trade of the Czech Republic (MPO ČR). (2021). *Action Plan for the National RIS3 Strategy: Vision and priorities for the 2021–2027 period*. https://ris3.gov.cz/en

Kautonen, T., Down, S., & Minniti, M. (2014). Ageing and entrepreneurial preferences. *Small Business Economics*, *42*(3), 579–594. https://doi.org/10.1007/s11187-013-9489-5

Kautonen, T., Down, S., & South, L. (2008). Enterprise support for older entrepreneurs: The case of PRIME in the UK. *International Journal of Entrepreneurial Behavior & Research*, *14*(2), 85–101. https://doi.org/10.1108/13552550810863071

Kurek, S., & Rachwał, T. (2011). Development of entrepreneurship in ageing populations of The European Union. *Procedia – Social and Behavioral Sciences*, *19*, 397–405. https://doi.org/10.1016/j.sbspro.2011.05.147

Lent, R. W., & Brown, S. D. (1996). Social cognitive approach to career development: An overview. *The Career Development Quarterly*, *44*(4), 310–321. https://doi.org/10.1002/j.2161-0045.1996.tb00448.x

Lévesque, M., & Minniti, M. (2006). The effect of aging on entrepreneurial behavior. *Journal of Business Venturing*, *21*(2), 177–194. https://doi.org/10.1016/j.jbusvent.2005.04.003

Maestripieri, D., & Boutwell, B. B. (2022). Human nature and personality variation: Reconnecting evolutionary psychology with the science of individual differences. *Neuroscience & Biobehavioral Reviews*, *143*, 104946. https://doi.org/10.1016/j.neubiorev.2022.104946

Martin, G. M., Austad, S. N., & Johnson, T. E. (1996). Genetic analysis of ageing: Role of oxidative damage and environmental stresses. *Nature Genetics*, *13*(1), 25–34. https://doi.org/10.1038/ng0596-25

Ministry for Regional Development (MMR). (2021). *Just Transition operational programme 2021–2027 (OP JT)*. Ministry for Regional Development (MMR). https://opst.cz/

Ministry of Education, Youth and Sports of the Czech Republic (MŠMT ČR). (2021). *National Research, Development and Innovation Strategy for Smart Specialisation of the Czech Republic (National RIS3 Strategy)*. https://mpo.gov.cz/en/business/ris3-strategy/

Mitchell, R. K., Busenitz, L., Lant, T., McDougall, P. P., Morse, E. A., & Smith, J. B. (2002). Toward a theory of entrepreneurial cognition: Rethinking the people side of entrepreneurship research. *Entrepreneurship Theory and Practice*, *27*(2), 93–104. https://doi.org/10.1111/1540-8520.00001

MPO. (2024). *Statistical data on entrepreneurs.* Ministry of Industry and Trade of the Czech Republic. https://mpo.gov.cz/en/business/licensed-trades/statistical-data-on-entrepreneurs/

Murnieks, C. Y., Mosakowski, E., & Cardon, M. S. (2014). Pathways of passion: Identity centrality, passion, and behavior among entrepreneurs. *Journal of Management*, *40*(6), 1583–1606. https://doi.org/10.1177/0149206311433855

Ness, R. K., Seifert, C. F., Marler, J. H., Wales, W. J., & Hughes, M. E. (2020). Proactive entrepreneurs: Who are they and how are they different? *The Journal of Entrepreneurship*, *29*(1), 148–175. https://doi.org/10.1177/0971355719893504

Oh, E., & Reeves, T. C. (2014). Generational differences and the integration of technology in learning, instruction, and performance. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 819–828). Springer. https://doi.org/10.1007/978-1-4614-3185-5_66

Perez-Encinas, A., Bueno, Y., Santos, B., & Nieto-Mejia, C. (2021). Are there differences and complementarities between senior and young entrepreneurs? An intergenerational perspective. *Sustainability*, *13*(9), 5202. https://doi.org/10.3390/su13095202

Rauch, A., & Frese, M. (2007). Let's put the person back into entrepreneurship research: A meta-analysis on the relationship between business owners' personality traits, business creation, and success. *European Journal of Work and Organizational Psychology*, *16*(4), 353–385. https://doi.org/10.1080/13594320701595438

Rubin, A., & Babbie, E. (1997). *Research methods for social work*. Brooks/Cole.

Seki, K., Ikeda, T., Urata, K., Shiratsuchi, H., Kamimoto, A., & Hagiwara, Y. (2022). Correlations between implant success rate and personality types in the older people: A preliminary case control study. *Journal of Dental Sciences*, *17*(3), 1266–1273. https://doi.org/10.1016/j.jds.2021.11.014

Shymko, Y., & Khoury, T. A. (2023). From community rootedness to individuated entrepreneuring: The development of entrepreneurial motivation through a temporary community of practice. *Journal of Business Venturing*, *38*(3), 106300. https://doi.org/10.1016/j.jbusvent.2023.106300

Stillman, D., & Stillman, J. (2017). *Gen Z @ work: How the next generation is transforming the workplace*. Harper Business.

Tooby, J., & Cosmides, L. (1990). On the universality of human nature and the uniqueness of the individual: The role of genetics and adaptation. *Journal of Personality*, *58*(1), 17–67. https://doi.org/10.1111/j.1467-6494.1990.tb00907.x

Trading Economics. (2023). *Self-employed, total (% of total employment) – Czech Republic*. https://tradingeconomics.com/

Waschkewitsch, L., & Kuhnt, A.-K. (2025). Babyboomer & Generation Z gemeinsam im Kreißsaal. *Hebammen Wissen*, *6*(1), 30–33 (In German). https://doi.org/10.1007/s43877-024-1255-1

Welpe, I. M., Spörrle, M., Grichnik, D., Michl, T., & Audretsch, D. B. (2012). Emotions and opportunities: The interplay of opportunity evaluation, fear, joy, and anger as antecedent of entrepreneurial exploitation. *Entrepreneurship Theory and Practice*, *36*(1), 69–96. https://doi.org/10.1111/j.1540-6520.2011.00481.x

Zachary, R. K., & Mishra, C. S. (2011). The future of enetrepreneurship research: Calling all researchers. *Entrepreneurship Research Journal*, *1*(1). https://doi.org/10.2202/2157-5665.1016

Zhang, T., & Acs, Z. (2018). Age and entrepreneurship: Nuances from entrepreneur types and generation effects. *Small Business Economics*, *51*(4), 773–809. https://doi.org/10.1007/s11187-018-0079-4

Zieger, B. (2024). Was Generationen voneinander lernen können. *Pflegezeitschrift*, *77*(8), 10–12 (In German). https://doi.org/10.1007/s41906-024-2657-1

The organizational innovation climate on affecting individuals' innovative work behavior: The roles of knowledge sharing and creative self-efficacy

Yu Li[1]*, Yang Chen*[2]*, Changfeng Wang*[3]

[1] Zhejiang Normal University, School of Economics and Management, Department of Business Administration, China, ORCID: 0009-0000-5955-8373, 860052600@qq.com;
[2] Zhejiang Yuexiu University, School of English Studies, Department of Translation and Interpreting, China, ORCID: 0000-0001-9024-3900, roy0901@163.com;
[3] Zhejiang Normal University, School of Economics and Management, Department of Business Administration, China, ORCID: 0000-0002-5236-4485, cfw@fudan.edu.cn (corresponding author).

Abstract: *Technological advancement promotes social changes and brings both opportunities and challenges to organizations. For high-tech enterprises in great demand of innovation and knowledge management, exploring what factors affect employees' innovative work behavior becomes emergent. This study investigates the impact of organizational innovation climate on employees' innovative work behavior, with knowledge sharing and creative self-efficacy as mediators. The structural model was tested using sample data from 307 employees of high-tech enterprises across different industries. The results demonstrated that the organizational innovation climate positively and significantly affected employees' innovative work behavior. Moreover, knowledge sharing and creative self-efficacy played chain mediation roles between organizational innovation climate and employees' innovative work behavior. At the end of the study, we propose practical implications for enterprises regarding organizational support, knowledge management, and incentive mechanism on individual innovation from internal and external sources.*

Keywords: *High-tech enterprises, organizational innovation climate, knowledge sharing, creative self-efficacy, innovative work behavior.*

JEL Classification: *J24, O31.*

APA Style Citation: Li, Y., Chen, Y., & Wang, C. F. (2026). The organizational innovation climate on affecting individuals' innovative work behavior: The roles of knowledge sharing and creative self-efficacy. *E&M Economics and Management, 29*(2), 61–75. https://doi.org/10.15240/tul/001/2026-5-001

Early Access Publication Date: *April 8, 2026.*

Introduction

In the current era, technological advancement such as the development of artificial intelligence has brought both opportunities and challenges to organizations. Many enterprises have sought to facilitate innovation performance with the adoption of technology to keep competitive (Li et al., 2023; Stacho et al., 2023). For high-tech enterprises featured by rapid updating of knowledge and dynamic technological transformation, the innovation is particularly important to their survival in the dynamic market.

As employees are valuable assets of organizations and their innovative work behavior (IWB) greatly affects enterprise innovation (Guo et al., 2023), exploring factors related to individual IWB becomes emergent.

Since innovation is a highly risky and demanding activity, individual IWB is affected by their perceived organization support, thus, the organizational innovation climate (OIC) has largely determined employees' innovation. OIC is a multifaceted construct that encompasses the employees' shared perceptions of the work environment including practices, policies, and procedures that facilitate or hinder innovation (Amabile & Gryskiewicz, 1989). It is a crucial element influencing employees' IWB in daily work, which creates conditions for the generation of original ideas and risk-taking initiatives (West & Anderson, 1996). When employees perceive the organizational support on their novel ideas or proposals, they are highly motivated to renovate at work, as individuals' innovativeness could be stimulated from the inspiring nature of climate (Liu et al., 2019). Previous research has reported the positive impact of OIC on employees' IWB (Alshahrani et al., 2025; Xu & Suntrayuth, 2022). However, as highlighted by a systematic review on innovation climate (Newman et al., 2019), the impact of OIC on IWB is dynamic and intricate when joining other contextual or individual factors. One possible explanation could be attributed to the mediating mechanisms on how OIC translates into individuals' IWB when different antecedents come into play. Therefore, there needs a holistic approach to address how and through which pathways OIC can effectively influence employees' IWB.

High-technology firms operate in environments characterized by short product lifecycles and intense competition, where knowledge sharing (KS) is crucial in enabling employees to build upon existing knowledge, generate new ideas, and develop innovative solutions to complex problems (Wang, 2025). During the KS process, employees' knowledge barriers may be removed, thus stimulating them to generate innovative ideas and apply them to handle real-world issues at workplace. Further, when organizations create an innovation-supportive climate that encourages KS among individuals, employees might gain the organization support and knowledge base for their innovativeness (Lu et al., 2021). Therefore, it could be indicated that KS plays an important role between OIC and IWB among those employees in high-tech enterprises, the context where innovation is both a strategic imperative and a survival requirement (Chughtai & Khalid, 2023; Mehmood et al., 2020).

As research continues, IWB is not only affected by organizational factors, but also determined by individual factors including personal characteristics, intrinsic motivation and self-efficacy (Srirahayu et al., 2023). From the individual aspect, creative self-efficacy (CSE) is one excellent internal predictor of IWB, as it relates to one's belief in the capacity for knowledge and expertise to generate creative outcomes (Jaiswal & Dhar, 2015). Employees with higher CSE take the initiative to acquire knowledge and have the self-confidence to confront challenges and setbacks, thereby performing persistently in carrying out innovative activities (Xu et al., 2021). Moreover, as the degree of organizational support perceived by employees significantly affects employees' CSE (Liu & Shi, 2009), this study hypothesizes that OIC might indirectly affect employees' IWB via their CSE in high-tech settings.

Despite the growing studies investigating employees' IWB as mentioned above, there remains a gap in understanding how the joint interplay of social and psychological processes translates OIC (the organizational factor) into individual IWB. To bridge the aforesaid research gap, this study investigates how OIC influences employees' IWB through the sequential mediating roles of KS and CSE in the high-stakes, fast-paced industry. Further, the findings may offer insightful references for entrepreneurs and human resource managers who wish to enhance their employees' innovation potentials and improve organizational innovation performance.

1 Theoretical background

This study is anchored in an integrated theoretical framework combining social exchange theory (SET) (Blau, 1964) and social cognitive theory (SCT) (Bandura, 1986), which collectively explain how OIC fosters IWB through sequential social and psychological mechanisms. SET is one of the most influential theories in the realm of management (Ren et al., 2025), including rules and norms of exchange, resources of exchange, and social exchange relationships (Ahmad et al., 2023). According to SET, employees reciprocate supportive organizational climates with prosocial behaviors (Cropanzano et al., 2017), such as knowledge sharing among co-workers, which are critical for innovation. Complementarily, SCT emphasizes the role of cognitive processes, particularly self-efficacy, in influencing individual behavior (Bandura, 1986). When employees share knowledge within a supportive innovation climate, they are exposed to diverse perspectives and expertise. This exposure subsequently bolsters their creative self-efficacy, as they gain confidence in their ability to generate novel ideas and solutions. With heightened CSE, employees are more likely to actively engage in innovative behavior, such as developing new products, improving processes, or suggesting creative solutions to problems. By integrating SET and SCT, our study provides a robust theoretical foundation for understanding the complex interplay among organizational factors, individual self-perceptions, and employees' innovative work behavior in the context of high-tech enterprises. This dual-lens framework justifies our proposed chain mediation model: OIC cultivates a culture of reciprocity (enhancing KS), which in turn bolsters employees' CSE that ultimately drives their IWB. The following section presents the relationship between the variables and then proposes the conceptual framework for this study.

1.1 Organizational innovation climate and innovative work behavior

OIC has been a hot topic due to the increasing significance of organizational innovation performance and employees' innovation attached to the fierce market competition. It is a multidimensional and complex concept that includes both objective physical environment and employees' subjective perceptions and feelings. The subjective perspective of OIC is widely accepted by scholars and it is positively associated with individuals' IWB (Newman et al., 2019). As innovation is a highly risky and demanding task, employees will have difficulties in generating novel ideas or renovating their work procedures without perceived organizational innovation support. In the earlier study, West and Anderson (1996) found that the supportive OIC could positively affect employees' endeavour to innovate in work-related practices. Recent literature also empirically supported a positive association between OIC and IWB; the higher the OIC perceived by employees, the stronger motivation they might acquire to innovate at work (Abderrahman Hassi et al., 2025; Güven & Doğan, 2025). As proposed by Xu and Suntrayuth (2022), a supportive OIC ensures that the necessary resources for innovation are available to employees, including time, funding, and access to information. Subsequently, with perceived innovation support within organisations and favourable resouces, employees develop the control of self-innovativeness and the intention to transform their novel ideas into innovative outcomes at work. For this study, OIC is hypothesized as an independent variable that predicts employees' IWB.

H1: Organizational innovation climate positively predicts employees' innovative work behavior.

1.2 Knowledge sharing as a mediator

Given the characteristics of rapid technological advancements and intense competition in high-tech enterprises, it is quite significant to encourage KS activities among knowledge workers to improve organizational performance (Souza & Paula, 2024). KS is a core activity in the realm of knowledge management, which emphasizes social interaction that may upgrade individuals' existing knowledge storage and contribute to the generation of novel ideas. However, KS cannot happen naturally without favorable conditions or incentive methods offered by organizations (Mehmood et al., 2020). Based on the reciprocate rule, when employees perceive that the organization values their creative ideas and provides a platform for innovation, they are more willing to share their tacit and explicit knowledge with their co-workers, which they believe will contribute to the overall innovative efforts of the organization. According to Zhang et al. (2025), organizational culture has a strong positive influence on knowledge sharing. For high-tech firms in particular, the more positive OIC perceived by employees, the more willing they are to share their work expertise and experiences with others, especially when it comes to tacit knowledge sharing (Lu et al., 2021). As highlighted by Xu and Suntrayuth (2022), employees in high-tech enterprises are more willing to share knowledge and solve technical difficulties with their colleagues when perceiving a supportive OIC. Consequently, knowledge sharing, with knowledge donation and knowledge collection occurring simultaneously, helps to rebuild individuals' existing knowledge framework, which positively influences employees' IWB at work (Wang, 2025). Therefore, the following hypothesis is proposed:

H2: Knowledge sharing mediates the relationship between organizational innovation climate and innovative work behavior.

1.3 Creative self-efficacy as a mediator

In the realm of business and management, CSE has been widely discussed and positively related to individual innovation and organizational performance (Gelaidan et al., 2023; Kafeel et al., 2023). CSE is a particular type of self-efficacy concerning one's self-confidence to generate creative outcomes. With internal confidence and belief, the knowledge workers' CSE could enhance their innovative behavior even when confronting uncertainties and challenges (Abidi et al., 2025). Empirical support for the predicted impact of CSE on employees' IWB was confirmed in different contexts, including R&D enterprises and educational institutions (e.g., Chen, 2024; Khan et al., 2023). Additionally, SCT posits that the person, environment, and behavior are interactive and influence each other (Wood & Bandura, 1989). In organizational contexts, a supportive OIC might allow employees to explore and experiment in their job-related activities, thereby gaining higher self-confidence in making innovative changes (Farmer & Tierney, 2017). With OIC triggering employees' CSE, employees gain self-concept and internal drives to actualize their innovativeness at work (Xu et al., 2021). Therefore, it is rational to propose the following hypothesis:

H3: Creative self-efficacy mediates the relationship between organizational innovation climate and innovative work behavior.

1.4 The chain intermediary effects of knowledge sharing and creative self-efficacy

As individuals' behavior is quite a complex process subject to both individual and environmental factors, it will be interesting to explore how different factors jointly affect employees' innovative work behavior at workplace. Based on the above explanations, the current study hypothesizes that when employees perceive a strong OIC, they are more willing to share their knowledge, expertise and work skills with others. Subsequently, employees upgrade their existing knowledge structure, thus increasing their self-confidence and proficiency in innovation (Islam et al., 2022). With considerable knowledge and equipped skills, employees boost their self-confidence in creatively dealing with challenging and risky situations (Khan et al., 2023), which contributes to their engagement in innovation at work. Therefore, we propose the following hypothesis and construct a conceptual model as Fig. 1.

H4: Knowledge sharing and creative self-efficacy play chain mediation roles between organizational innovation climate and innovative work behavior.

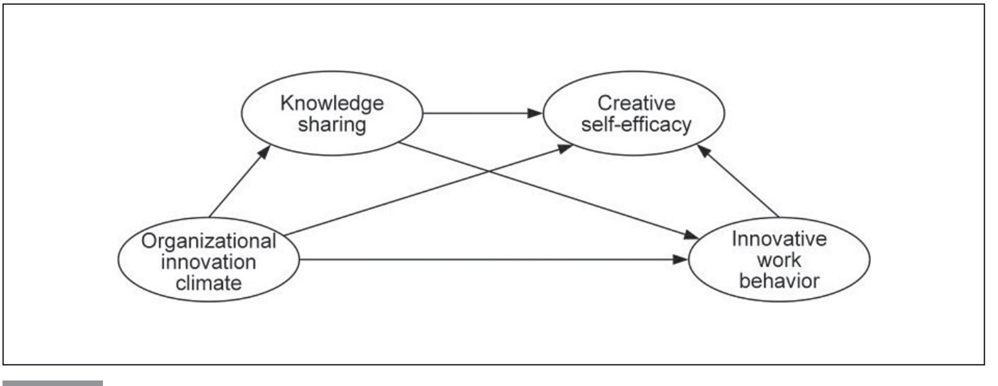

Fig. 1: Conceptual framework

Source: own

2 Research methodology

2.1 Sampling and data collection

This study used the convenience sampling method to recruit employees from high-tech enterprises in Zhejiang Province, the southeastern area of China. The reason for selecting this population lies in the rapid development of high technology and economy supported by government in this area. Before administering the questionnaires, participants were informed of the research objectives, and they understood that their personal information was kept anonymous and confidential. All participants were voluntary to fill in the survey and they were in contact through the organization leaders or human resource managers. From April to May in 2023, the first researcher sent 335 questionnaires to the target participants. To avoid potential bias in the research design, the present study collected data twice, with an interval period of 2 weeks. In time 1, participants were asked to fill in their demographic information, their perceived OIC and IWB; then after two weeks, the same participants were required to fill in the questionnaires inquiring their KS behavior and CSE at work. Finally, 307 valid questionnaires were obtained, with a recovery rate of 91.64%. Tab. 1 presents the demographic information of the participants, including the gender, age, education levels, tenure and job position.

Tab. 1: Demographic information of participants

Variable	Category	Number	Percentage (%)
Gender	Male	180	58.6
	Female	127	41.4
Age	Below 25 years	26	8.5
	26–35 years	161	52.4
	36–45 years	98	31.9
	Above 46 years	22	7.2
Education	Junior college below	21	6.8
	Junior college	63	20.5
	Bachelor	148	48.2
	Master and over	75	24.4
Tenure	Below 2 years	81	26.4
	2–8 years	136	44.3
	8–15 years	71	23.1
	Above 15 years	19	6.2
Job position	Staff	195	63.5
	Junior manager	66	21.5
	Middle manager	30	9.8
	Top manager	16	5.2

Source: own

2.2 Measures and variable operation

All instruments of the present study were adjusted from previous studies and all items were measured on a seven-point Likert scale ranging from 1 to 7, with 1 representing “strongly disagree” to 7 representing “strongly agree”. The questionnaire consists of two sections, the first section was to obtain participants’ demographic information and the second included these variables, independent variable (OIC), mediators (KS and CSE), dependent variable (IWB). The following section reported the formal questionnaires addressed to participants.

Independent variable. The Organizational Innovation Climate Scale developed by Liu and Shi (2009) was employed to assess OIC, the independent variable of this study. This 12-item instrument consisted of three constructs, namely, team support (TS), supervisor support (SS) and organization support (OS). Sample items included, "*In our team, we often communicate and discuss on problems at work*" (TS); "*My supervisor respects different ideas proposed by their sub-ordinates*" (SS); "*My organization values openness and innovation*" (OS). The values of Cronbach's alpha coefficient of each construct were assessed to be 0.883, 0.890 and 0.887, respectively.

Mediating variables. Knowledge sharing was assessed by a five-item instrument, which was adjusted by Lu et al. (2006), with the value of Cronbach's alpha coefficient being 0.883. Sample items were, "*I share useful work experience with others*"; "*When acquiring new knowledge useful to work, I am willing to share with more people.*"

Creative self-efficacy was measured by a five-item instrument adjusted from Karwowski et al. (2018). Sample items included, "*I know I can efficiently solve complex problems*"; "*I trust my creative ability*." The value of Cronbach's alpha coefficient of this measurement was 0.891.

Dependent variable. The five-item innovative work behavior scale was adapted from Liu and Shi (2009). Sample items were, "*I often come up innovative ideas at work*"; "*I often share my innovative ideas with my colleagues or supervisors, to obtain their support*." The value of Cronbach's alpha coefficient of this measurement was 0.891.

Control variables. As suggested by previous research (Newman et al., 2018), this study controlled demographic variables in the analysis of our model. Due to the differences in social division of labor and workplace advantages, gender can influence employees' IWB. Therefore, this study considered gender as one of the control variables, which was coded as 0 = female, 1 = male. In addition, the education level was closely related to employees' cognitive levels and knowledge structures, which in turn affected their IWB (coded as 1 = junior college below; 2 = junior college; 3 = bachelor; 4 = master and above). Finally, the job position was regarded as another control variable and coded ordinally (1 = staff; 2 = junior managers; 3 = middle managers; 4 = top managers).

2.3 Descriptive statistics and correlation matrix

Descriptive statistics were calculated to summarize the sample characteristics (e.g., gender, education levels, job position) and the distribution of all study variables. The means, standard deviations, and correlations were presented in Tab. 2. The significant and positive correlations between the independent

Tab. 2: Descriptive statistics and correlation matrix

Variable	Mean	SD	1	2	3	4	5	6	7
1. Gender	0.586	0.493	1						
2. Education levels	2.902	0.846	0.059	1					
3. Job position	1.567	0.870	0.053	0.169**	1				
4. OIC	5.095	0.668	0.118*	0.190**	0.177**	1			
5. KS	4.775	0.750	0.152**	0.192**	0.162**	0.474**	1		
6. CSE	4.788	0.722	0.117*	0.179**	0.194**	0.428**	0.540**	1	
7. IWB	4.496	0.717	0.135*	0.177**	0.180**	0.510**	0.525**	0.612**	1

Note: * $p < 0.05$; ** $p < 0.01$; OIC – organizational innovation climate; KS – knowledge sharing; CSE – creative self-efficacy; IWB – innovative work behavior; SD – standard deviation.

Source: own

variable, mediators and dependent variable could provide preliminary support for further hypotheses testing.

2.4 Data analysis

We conducted the hierarchical linear modeling to test the hypotheses and the mediation analysis was performed using the software Mplus 8.3, with the 5,000 bootstrapping method. The mediation effect was considered significant if the 95% confidence interval did not include zero.

Prior to testing the hypotheses, the reliability and validity of the variables were required to be assessed. According to Fornell and Larcker (1981), the reliability was considered acceptable if the values of Cronbach's alpha coefficient and the composite reliability (CR) coefficient were above 0.7. The convergent validity was assessed by extracted mean variance (AVE), and the discriminant validity could be assessed by analyzing the relationship between the correlation coefficient of each variable and the square root of AVE. We also conducted confirmatory factor analysis (CFA) to test the measurement model, following these goodness-of-fit indices: χ^2/df (≤3.0), comparative fit index (CFI ≥ 0.90), Tucker-Lewis index (TLI ≥ 0.90), root mean square error of approximation (RMSEA ≤ 0.08), standardized root mean square residual (SRMR ≤ 0.08) (Hair et al., 2010). Additionally, Akaike information criterion (AIC) and Bayesian information criterion (BIC) were reported, with lower values indicating better model fit.

3 Research results

3.1 Reliability and validity analysis

Tab. 3 presents the results of reliability, convergent and discriminant validity among the variables. For this study, the reliability was considered to be good, as the composite reliability (CR) of all variables and factor loadings of each items were above the recommended value of 0.7. Additionally, Tab. 3 demonstrated that the values of AVE of each construct were all above 0.5, showing good convergence validity of measurements as suggested by Bagozzi and Yi (1988). Lastly, the square root value of AVE of all variables was more significant than the correlation coefficient between variables, showing good discriminant validity.

We also employed Harman's on-factor method to check the common method bias (CMB), providing evidence that the first factors in the model explained 39.4% of the variance, which was less than the recommended 50% (Podsakoff et al., 2003). Therefore, CMB was not a potential problem in this study.

Tab. 3: Reliability, convergent and discriminant validity among variables

Variable	Loadings	CR	AVE	1	2	3	4	5	6
1. TS	0.804–0.815	0.883	0.654	**0.809**					
2. SS	0.799–0.839	0.890	0.669	0.631***	**0.818**				
3. OS	0.808–0.818	0.887	0.662	0.622***	0.610***	**0.814**			
4. KS	0.758–0.794	0.883	0.601	0.461***	0.460***	0.428***	**0.775**		
5. CSE	0.767–0.808	0.891	0.621	0.433***	0.364***	0.416***	0.612***	**0.788**	
6. IWB	0.765–0.803	0.891	0.620	0.429***	0.487***	0.517***	0.589***	0.689***	**0.787**

Note:****p* < 0.001; TS – team support; SS – supervisor support; OS – organization support; KS – knowledge sharing; CSE – creative self-efficacy; IWB – innovative work behavior.

Source: own

3.2 Measurement model

This study conducted CFA to verify the fitness of the proposed model and Tab. 4 presents the results. As suggested by Hair et al., (2010), our six-factor measurement model fit the data well (χ^2/df = 1.254, AIC = 525.465, BIC = 782.617, RMSEA = 0.029, TLI = 0.982, CFI = 0.984, SRMR = 0.032), which demonstrated to be superior to the alternative four-factor, two-factor and one-factor models. Therefore, these good fit values provided the prerequisite condition for our hypotheses testing.

Tab. 4: Results of confirmatory factor analysis (*N* = 307)

Model	χ^2	df	AIC	BIC	RMSEA	TLI	CFI	SRMR
1. Six-factor model	387.465	309	525.465	782.617	0.029	0.982	0.984	0.032
2. Four-factor model (TS + SS + OS, KS, CSE, IWB)	1,006.115	318	1,126.115	1,349.726	0.084	0.850	0.864	0.058
3. Two-factor model (TS + SS + OS + KS, CSE + IWB)	1,851.614	323	1,961.614	2,166.591	0.124	0.671	0.697	0.103
4. One-factor model	2,430.168	324	2,430.168	2,430.168	0.146	0.548	0.583	0.118

Note:*** $p < 0.001$; TS – team support; SS – supervisor support; OS – organization support; KS – knowledge sharing; CSE – creative self-efficacy; IWB – innovative work behavior.

Source: own

3.3 Hypotheses testing

This study conducted a hierarchical linear modeling to test the hypotheses. We first entered all three control variables into a baseline model and the regression result showed that the gender significantly affected employees' IWB ($\beta = 0.172$, $p < 0.05$). This result indicated that the male employees demonstrated significantly higher IWB than females. Additionally, the beta estimate for education levels ($\beta = 0.123$, $p < 0.05$) and job positions were positive ($\beta = 0.123$, $p < 0.01$), indicating that employees with higher education levels reported greater IWB, and higher-ranking positions demonstrated significantly stronger IWB. Next, OIC was added as the independent variable and the regression result indicated that OIC positively predicted employees' IWB ($\beta = 0.509$, $p < 0.001$), which supported *H1* for this study.

The study adopted the bootstrapping method to test the mediation effects, and the data fit our model well with χ^2/df =1.232, GFI = 0.909, AGFI = 0.890, TLI = 0.980, CFI = 0.982, SRMR = 0.035, RMSEA = 0.028. Fig. 2 presents the chain mediating model between OIC and IWB, with KS and CSE as the mediators. Tab. 5 showed the results of 5,000 bootstrap samples, and there was no zero value in the 95% confidence interval. Seeing from Tab. 5, Ind1 showed that KS positively mediated the relationship between OIC and IWB (b = 0.093, SE = 0.048, Z = 1.968, $p < 0.05$, LL/UL = 0.011/0.202), providing support for *H2*. Also, CSE positively mediated the relationship between OIC and IWB (b = 0.121, SE = 0.051, Z = 2.366, $p < 0.05$, LL/UL = 0.0029/0.232), providing support for *H3*. Meanwhile, KS and CSE positively and significantly played the chain mediating role between OIC and IWB (b = 0.132, SE = 0.037, Z = 3.605, $p < 0.001$, LL/UL = 0.076/0.227). Therefore, *H4* was supported in this study.

Tab. 5: Mediation analysis results

	Effect	SE	Z	*p*	LL	UL
Direct effect	0.348	0.096	3.628	***	0.179	0.552
Total indirect effect	0.347	0.071	4.882	***	0.220	0.504
OIC→KS→IWB (Ind1)	0.093	0.048	1.968	*	0.011	0.202
OIC→CSE→IWB (Ind2)	0.121	0.051	2.366	*	0.029	0.232
OIC→KS→CSE→IWB (Ind3)	0.132	0.037	3.605	***	0.076	0.227

Note:* $p < 0.05$; *** $p < 0.001$; OIC – organization innovation climate: KS – knowledge sharing; CSE – creative self-efficacy; IWB – innovative work behavior; SE – standard error; LL – lower confidence level; UL – upper confidence level; Bootstrap sample size = 5,000; control variables added.

Source: own

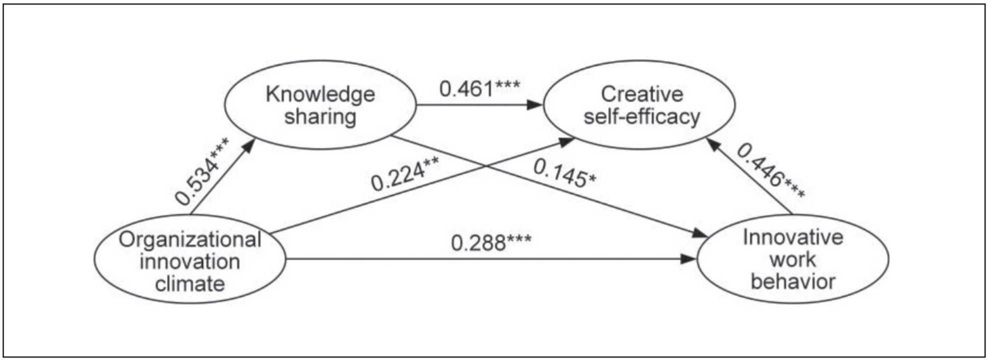

Fig. 2: **SEM results**

Source: own

4 Discussion

Innovative work behavior is a complex process that emphasizes the willingness and efforts of individuals to innovate, relying on both individual capabilities, motivations and external resources (Zhu et al., 2022). This study examined how OIC influences employees' IWB through the chain mediation of KS and CSE in high-tech industries. The findings confirm all hypotheses, offering nuanced insights into the mechanisms linking contextual, social, and individual motivational factors to innovation. Below, we discuss the major findings, theoretical and practical implications and outline avenues for future research.

First, the findings demonstrate that OIC positively predict employees' IWB, which is consistent with previous studies (Abderrahman Hassi et al., 2025; Xu & Suntrayuth, 2022) reporting that a positive OIC empowered employees to renovate their work procedures or practices. A supportive innovation climate within organizations, characterized by management endorsement, resource availability, and freedom to experiment, creates a psychological safety net for employees in the high-tech industry. This environment encourages the generation of novel ideas, risk-taking initiatives and the pursuit of creative solutions, which are stimulus of innovative behavior. Conversely, when employees perceive unfavorable innovation support from organizations, their motivation and enthusiasm to innovate are hampered for fear of making mistakes and being blamed (Cai & Tang, 2021). More importantly, compared to studies concluding that innovation is likely to be inspired in the western culture that values individualism (e.g., Taylor & Wilson, 2012), our study enriches the mechanism of OIC on employees' IWB in the Asian context where collectivism prevails.

Unlike previous research that predominantly focused on the independent mediating roles of these factors (Azeem et al., 2021; Teng et al., 2019), we reveal a sequential pathway. This study contributes to exploring the influence process of OIC on employees' IWB by joining KS and CSE as chain mediators. On one hand, the findings confirm the mediating role of KS between OIC and IWB, which is in align with a recent systematic review concluding that the organizational support and reciprocal relationships promote employees' KS activities within organizations, which in turn inspire employees' innovation and creativity (Santhose & Lawrence, 2023). In the context of this research, a positive organizational innovation climate that encourages risk-taking initiatives and offers support to employees, represents the "rewards" provided by the organization. In response, employees engage in KS as a form of reciprocation, believing that their contributions will lead to personal benefits such as recognition or career advancement. On the other hand, our study extends previous studies reporting KS behaviors strengthen individuals' CSE (Islam & Asad, 2024; Khan et al., 2023), which is critical for managing uncertainty in innovation commitments. In the notion of SCT, the role of self-efficacy is highlighted in driving individuals to undertake challenging tasks with persistence in facing obstacles or failures. When employees engage in sharing knowledge with their co-workers, they gain exposure to diverse perspectives and expertise, which bolsters their confidence in their creative capabilities. This increased creative self-efficacy, in turn, serves as a powerful motivator for employees to engage in innovative activities, forming a chain reaction from organizational climate to innovative behavior. Connecting factors from both external resources and individual perspectives, this study incorporates OIC, KS, CSE and IWB into one single model, revealing different impact mechanisms on influencing employees' innovative work behavior.

Theoretically, this study makes significant contributions by integrating SET and SCT to explore the intricate relationship between OIC, KS, CSE, and employees' IWB. Prior studies have predominantly examined these theories in isolation, which may limit the comprehensiveness of understanding complex behavioral processes. Our findings fill in this gap by integrating both theories, demonstrating that the effects of OIC are not merely contextual or psychological but arise from their synergy – a theoretical leap underscored by recent calls for investigating multi-faceted antecedents related to innovative work behavior (Alessa & Durugbo, 2022). To our knowledge, the impact of OIC on employees' IWB via the intermediary effects of KS and CSE has not been explored previously. Our study provides a novel theoretical perspective, illustrating how organizational-level factors can influence individual-level innovative behavior through both social exchange and cognitive processes. It contributes to the development of a more integrated and explanatory theory in the field of organizational innovation, offering scholars a new analytical framework to explore the complex dynamics of innovation-related behaviors.

4.1 Practical implications

This empirical study was conducted in the settings of high-tech enterprises, and our findings might provide some practical implications for organizations that intend to enhance organizational innovation outcomes. First, as OIC positively affects employees' IWB, organizations are suggested to create an innovation-supportive environment, which might inspire employees' motivation and enthusiasm to engage in innovative activities. For high-tech enterprises in particular, those knowledge-based employees are in high demand for certain job autonomy and flexibility regarding work arrangements. Therefore, it is crucial to provide employees with sufficient respect and trust and be tolerant of the possible failure resulting from employees' innovative attempts at work. It is suggested that supervisors establish an equal and friendly relationship with their subordinates and provide positive feedback to employees' innovative commitments (Su & Zhang, 2022). Managers could also attend target training programs to role model their IWB (Newman et al., 2019) and guide the team to support their peers' innovativeness. In this way, when employees perceive a friendly innovative climate within organizations, they may develop the willingness to express their novel ideas and gain psychological safety in carrying out innovative activities.

Second, organizations need to offer a convenient knowledge-sharing platform to employees with certain technical and tool support (Xu & Suntrayuth, 2022). For knowledge-extensive enterprises, their employees need to acquire new knowledge and skills through information exchange, meanwhile the collision of diverse ideas enhances employees' engagement in innovation. Nevertheless, it requires managers' attention that employees may seek to hide their know-how and expertise for fear of losing their competitive advantages (Lu et al., 2012). In this regard, organizations need to provide positive feedback and substantial rewards for employees' KS behavior so that employees feel fair treatment and organizational justice (Akram et al., 2019). According to SET, individuals seek reciprocity from those who provide them with benefits in return. Only based on mutual benefit might increase employees' willingness to share their expertise with others, further stimulating them to engage in collective innovation activities.

Finally, considering the role of CSE in triggering employees' IWB, the human resource department should consider the person-job fit factor in matching the knowledge, personality or skills of employees with their daily work arrangements. When employees feel interested and competent in their assigned tasks, they will gain the sense of fulfilment that internally triggers their CSE in innovative commitments. It is also advisable to adopt CSE intervention designs and implement CSE enhancement programs on employees within organizations (Farmer & Tierney, 2017). Moreover, considering the chain mediating roles of KS and CSE in the relationship between OIC and IWB, it is necessary to adopt a range of incentives that enhance employees' readiness and willingness to engage in innovative activities. For high-tech enterprises characterized by rapid upgrading of knowledge and technology, providing employees with target training programs enables them to acquire the latest knowledge and generate creative ideas. With sufficient knowledge, those employees may feel competent with determination and self-confidence (indicating higher CSE) even in the risky and challenging innovation attempts. As transferring from creative ideas into practical outcomes requires time and resources, funding support and trust on employees could reduce the psychological pressure that may hinder their innovation performance.

Conclusions

This study aimed to explore the intricate processes by which organizational innovation climate affects employees' innovative work behavior in the high-tech industry. Theoretically, this study contributes to offering a novel perspective in understanding how the organizational factor transfers into individual behavior, by uncovering the chain mediation mechanism of knowledge sharing and creative self-efficacy. Practically, managers are recommended to prioritize cultivating a positive innovation climate by providing sufficient resources to encourage the generation of novel ideas at work. It is necessary to promote communication, and rewarding innovation to encourage knowledge sharing, while also implementing strategies like training programs and mentorship to boost employees' creative self-efficacy, thereby enhancing the overall innovation capabilities of enterprises.

However, this study is not without limitations. First, the data collection was restricted to the high-tech industry, which may limit the generalizability of the research results. Since different industries have unique characteristics and work cultures, and the relationships explored in this study may vary across industries. Future research could expand the sample to include multiple industries, such as the service industry and the manufacturing industry, to comprehensively test the robustness of the proposed model. Second, since employees' innovative work behavior is a dynamic and complex process, which needs to be observed and assessed over a long period. As this is a quantitative study employing cross-sectional data to analyze factors influencing employees' innovative work behavior, future researchers can adopt qualitative research methods such as interviews and field observations to gain a more in-depth understanding of factors influencing employees' innovative work behaviors. Additionally, longitudinal studies are needed to track the dynamic changes over time, which can provide a more in-depth understanding of how organizational innovation climate, knowledge sharing, creative self-efficacy, and innovative behavior interact and evolve. Finally, only two mediating variables were considered in this study. Future research are suggested to investigate other potential mediating or moderating variables, such as leadership style, organizational culture at a deeper level, and individual personality traits. These additional factors may further enrich the understanding of the complex mechanisms underlying employees' innovative behavior.

In conclusion, this study has made important progress in exploring the influence mechanism of organizational innovation climate on employees' innovative work behavior. The identification of the chain mediation roles of knowledge sharing and creative self-efficacy provides a new theoretical and practical roadmap for understanding the innovation process in high-tech organizations. While the existing limitations point to directions for future research, the findings of this study have laid a solid foundation for both academic research and practical management in promoting organizational innovation.

Acknowledgments: *Supported by The National Social Science Fund of China (22AGL006).*

References

Abderrahman Hassi, A., Rohlfer, S., & Jebsen, S. (2025). Organizational climate for initiative and innovative work behavior: A moderated mediation model. *Evidence-Based HRM: A Global Forum for Empirical Scholarship*, *14*(1), 192–206. https://doi.org/10.1108/ebhrm-03-2024-0097

Abidi, O., Safi, M., Sarhan, H., & Bhagat, R. (2025). Creative self-efficacy and innovative behaviour amidst the digital deluge: Unravelling their nexus with autonomy and workplace civility. *International Journal of Productivity and Performance Management*, *74*(8), 2883–2907. https://doi.org/10.1108/ijppm-05-2024-0319

Ahmad, R., Nawaz, M. R., Ishaq, M. I., Khan, M. M., & Ashraf, H. A. (2023). Social exchange theory: Systematic review and future directions. *Frontiers in Psychology*, *13*(2), 1–13. https://doi.org/10.3389/fpsyg.2022.1015921

Akram, T., Lei, S., Haider, M. J., & Hussain, S. T. (2019). The impact of organizational justice on employee innovative work behavior: Mediating role of knowledge sharing. *Journal of Innovation & Knowledge*, *5*(2), 117–129. https://doi.org/10.1016/j.jik.2019.10.001

Alessa, H. S., & Durugbo, C. M. (2022). Systematic review of innovative work behavior concepts and contributions. *Management Review Quarterly*, *72*(4), 1171–1208. https://doi.org/10.1007/s11301-021-00224-x

Alshahrani, M. A., Yaqub, M. Z., Ali, M., El Hakimi, I., & Salam, M. A. (2025). Could entrepreneurial leadership promote employees' IWB? The roles of intrinsic motivation, creative self-efficacy and firms' innovation climate. *International Journal of Innovation Science* (in press). https://doi.org/10.1108/ijis-08-2024-0211

Amabile, T. M., & Gryskiewicz, N. D. (1989). The creative environment scales: Work environment inventory. *Creativity Research Journal*, *2*(4), 231–253. https://doi.org/10.1080/10400418909534321

Azeem, M., Ahmed, M., Haider, S., & Sajjad, M. (2021). Expanding competitive advantage through organizational culture, knowledge sharing and organizational innovation. *Technology in Society*, *66*(1), 101635. https://doi.org/10.1016/j.techsoc.2021.101635

Bagozzi, R. P., & Yi, Y. (1988). On the evaluation of structural equation models. *Journal of the Academy of Marketing Science*, *16*(1), 74–94. https://doi.org/10.1007/bf02723327

Bandura, A. (1986). *Social foundations of thought and action*. Prentice Hall.

Blau, P. M. (1964). *Exchange and power in social life*. John Wiley and Sohns.

Cai, Y., & Tang, R. (2021). School support for teacher innovation: Mediating effects of teacher self-efficacy and moderating effects of trust. *Thinking Skills and Creativity*, *41*, 100854. https://doi.org/10.1016/j.tsc.2021.100854

Chen, Y. (2024). Teacher autonomy, creative self-efficacy, and innovative behavior: Perspectives from Chinese university EFL teachers. *Arab World English Journal*, *15*(2), 73–86. https://doi.org/10.24093/awej/vol15no2.5

Chughtai, M. S., & Khalid, Y. (2023). Learning organizations and innovative work behaviors: A moderated mediation model from the perspective of social schema theory. *Organizational Psychology*, *13*(2), 207–229. https://doi.org/10.17323/2312-5942-2023-13-2-207-229

Cropanzano, R., Anthony, E. L., Daniels, S. R., & Hall, A. V. (2017). Social exchange theory: A critical review with theoretical remedies. *Academy of Management Annals*, *11*(1), 479–516. https://doi.org/10.5465/annals.2015.0099

Farmer, S. M., & Tierney, P. (2017). Considering creative self-efficacy: Its current state and ideas for future inquiry. In M. Karwowski & J. C. Kaufman (Eds), *The creative self* (pp. 23–47). Elsevier. https://doi.org/10.1016/b978-0-12-809790-8.00002-9

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, *18*(1), 39–50. https://doi.org/10.2307/3151312

Gelaidan, H. M., Al-Swidi, A. K., & Al-Hakimi, M. A. (2023). Servant and authentic leadership as drivers of innovative work behaviour: The moderating role of creative self-efficacy. *European Journal of Innovation Management*, *27*(6), 1938–1966. https://doi.org/10.1108/ejim-07-2022-0382

Guo, Y., Peng, Y., & Zhu, Y. (2023). How does empowering leadership motivate employee innovative behavior: A job characteristics perspective. *Current Psychology*, *42*(21), 18280–18290. https://doi.org/10.1007/s12144-022-03000-6

Güven, Ş., & Doğan, E. (2025). The impact of organizational climate on innovative behavior: The mediating role of work engagement. *Ekoist: Journal of Econometrics and Statistics*, *42*(42), 209–222. https://doi.org/10.26650/ekoist.2025.42.1644637

Hair, J., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis: A global perspective* (7th ed.). Pearson Education.

Islam, T., & Asad, M. (2024). Enhancing employees' creativity through entrepreneurial leadership: Can knowledge sharing and creative self-efficacy matter? *VINE Journal of Information and Knowledge Management Systems*, *54*(1), 59–73. https://doi.org/10.1108/vjikms-07-2021-0121

Islam, T., Zahra, I., Rehman, S. U., & Jamil, S. (2022). How knowledge sharing encourages innovative work behavior through occupational self-efficacy? The moderating role of entrepreneurial leadership. *Global Knowledge, Memory and Communication*, *73*(1/2), 67–83. https://doi.org/10.1108/gkmc-02-2022-0041

Jaiswal, N. K., & Dhar, R. L. (2015). Transformational leadership, innovation climate, creative self-efficacy and employee creativity: A multilevel study. *International Journal of Hospitality Management*, *51*, 30–41. https://doi.org/10.1016/j.ijhm.2015.07.002

Kafeel, L., Khan, M. M., & Ahmed, S. S. (2023). Authentically flown to innovation: Authentically triggered innovative work behavior through flow at work with moderating role of creative self-efficacy. *International Journal of Innovation Science*, *16*(3), 463–481. https://doi.org/10.1108/ijis-05-2022-0099

Karwowski, M., Lebuda, I., & Wiśniewska, E. (2018). Measuring creative self-efficacy and creative personal identity. *The International Journal of Creativity and Problem Solving*, *28*(1), 45–57.

Khan, H. S. ud din, Li, P., Chughtai, M. S., Mushtaq, M. T., & Zeng, X. (2023). The role of knowledge sharing and creative self-efficacy on the self-leadership and innovative work behavior relationship. *Journal of Innovation & Knowledge*, *8*(4), 100441. https://doi.org/10.1016/j.jik.2023.100441

Li, S., Gao, L., Han, C., Gupta, B., Alhalabi, W., & Almakdi, S. (2023). Exploring the effect of digital transformation on firms' innovation performance. *Journal of Innovation & Knowledge*, *8*(1), 100317. https://doi.org/10.1016/j.jik.2023.100317

Liu, F., Chow, I. H.-S., Zhang, J.-C., & Huang, M. (2019). Organizational innovation climate and individual innovative work behavior: Exploring the moderating effects of psychological ownership and psychological empowerment. *Review of Managerial Science*, *13*(4), 771–789. https://doi.org/10.1007/s11846-017-0263-y

Liu, Y., & Shi, J. (2009). A study on the relationship between the effects of the organizational innovative climate and those of motivational preference, on employees' innovative work behavior. *Management World*, *193*(10) (in Chinese). https://doi.org/10.19744/j.cnki.11-1235/f.2009.10.011

Lu, L., Leung, K., & Koch, P. T. (2006). Managerial knowledge sharing: The role of individual, interpersonal, and organizational factors. *Management and Organization Review*, *2*(1), 15–41. https://doi.org/10.1111/j.1740-8784.2006.00029.x

Lu, L., Lin, X., & Leung, K. (2012). Goal orientation and innovative performance: The mediating roles of knowledge sharing and perceived autonomy. *Journal of Applied Social Psychology*, *42*(S1), 180–197. https://doi.org/10.1111/j.1559-1816.2012.01018.x

Lu, X., Huang, J., & Song, J. (2021). Effect of innovation atmosphere on innovation performance in Chinese and South Korean high-tech enterprises: Mediating role of tacit knowledge sharing. *Science and Technology Management Research*, *10*. https://doi.org/10.3969/j.issn.1000-7695.2021.10.020

Mehmood, M. S., Jian, Z., & Akram, U. (2020). Be so creative they can't ignore you! How can entrepreneurial leader enhance the employee creativity? *Thinking Skills and Creativity*, *38*, 100721. https://doi.org/10.1016/j.tsc.2020.100721

Newman, A., Round, H., Wang, S., & Mount, M. (2019). Innovation climate: A systematic review of the literature and agenda for future research. *Journal of Occupational and Organizational Psychology*, *93*(1), 73–109. https://doi.org/10.1111/joop.12283

Newman, A., Tse, H. H. M., Schwarz, G., & Nielsen, I. (2018). The effects of employees' creative self-efficacy on innovative behavior: The role of entrepreneurial leadership. *Journal of Business Research*, *89*(8), 1–9. https://doi.org/10.1016/j.jbusres.2018.04.001

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

Ren, Q., Xu, X., Scullion, R., Lin, Y., Wang, S., & Rong, K. (2025). Linking social exchange theory to B2B relationship innovation management: The moderating role of reciprocity. *Technological Forecasting and Social*

Change, *212*, 124003. https://doi.org/10.1016/j.techfore.2025.124003

Santhose, S. S., & Lawrence, L. N. (2023). Understanding the implementations and limitations in knowledge management and knowledge sharing using a systematic literature review. *Current Psychology*, *42*(36), 32427–32442. https://doi.org/10.1007/s12144-022-04115-6

Souza, F., & Paula, F. (2024). From authentic leadership to knowledge sharing and knowledge worker retention: Mediating role of psychological empowerment. *Knowledge Management Research & Practice*, *23*(5), 493–506. https://doi.org/10.1080/14778238.2024.2400249

Srirahayu, D. P., Ekowati, D., & Sridadi, A. R. (2023). Innovative work behavior in public organizations: A systematic literature review. *Heliyon*, *9*(2), e13557. https://doi.org/10.1016/j.heliyon.2023.e13557

Stacho, Z., Stachová, K., Lukáč, M., Kupec, V., & Petrů, N. (2023). Management reflections on innovations in digitalization, with an emphasis on degree of work autonomy. *E&M Economics and Management*, *26*(1), 78–93. https://doi.org/10.15240/tul/001/2023-1-005

Su, W., & Zhang, Y. (2022). More positive, more innovative: A moderated-mediation model of supervisor positive feedback and subordinate innovative behavior. *Current Psychology*, *42*(33), 29682–29694. https://doi.org/10.1007/s12144-022-04047-1

Taylor, M. Z., & Wilson, S. (2012). Does culture still matter?: The effects of individualism on national innovation rates. *Journal of Business Venturing*, *27*(2), 234–247. https://doi.org/10.1016/j.jbusvent.2010.10.001

Teng, C., Hu, C., & Chang, J. (2019). Triggering creative self-efficacy to increase employee innovation behavior in the hospitality workplace. *The Journal of Creative Behavior*, *54*(4), 912–925. https://doi.org/10.1002/jocb.419

Wang, Z. (2025). Linking innovative knowledge sharing and employees' innovative behaviour: The mediating role of thriving at work. *Knowledge Management Research & Practice*, *23*(1), 52–62. https://doi.org/10.1080/14778238.2023.2261411

West, M. A., & Anderson, N. R. (1996). Innovation in top management teams. *Journal of Applied Psychology*, *81*(6), 680–693. https://doi.org/10.1037//0021-9010.81.6.680

Wood, R., & Bandura, A. (1989). Social cognitive theory of organizational management. *The Academy of Management Review*, *14*(3), 361–384. https://doi.org/10.2307/258173

Xu, H., Guo, P., & Bao, L. (2021). Influence of organizational innovation support on innovative work behavior of researchers: Chain intermediary effect of creative self-efficacy and knowledge sharing. *Science and Technology Management Research*, *8*, 124–131. https://doi.org/10.3969/j.issn.1000-7695.2021.08.017

Xu, Z., & Suntrayuth, S. (2022). Innovative work behavior in high-tech enterprises: Chain intermediary effect of psychological safety and knowledge sharing. *Frontiers in Psychology*, *13*, 1–11. https://doi.org/10.3389/fpsyg.2022.1017121

Zhang, J., Jehangir, F. N., Yang, L., Tahir, M. A., & Tabasum, S. (2025). Competitive advantage and firm performance: The role of organizational culture, organizational innovation, and knowledge sharing. *Journal of the Knowledge Economy*, *16*(1), 3081–3107. https://doi.org/10.1007/s13132-024-01910-3

Zhu, Y., Liu, J., Lin, S., & Liang, K. (2022). Unlock the potential of regional innovation environment: The promotion of innovative behavior from the career perspective. *Journal of Innovation & Knowledge*, *7*(3), 100206. https://doi.org/10.1016/j.jik.2022.100206

West, M. A., & Anderson, N. R. (1996). Innovation in top management teams. *Journal of Applied Psychology*, *81*(6), 680–693. https://doi.org/10.1037//0021-9010.81.6.680

Wood, R., & Bandura, A. (1989). Social cognitive theory of organizational management. *The Academy of Management Review*, *14*(3), 361–384. https://doi.org/10.2307/258173

Xu, H., Guo, P., & Bao, L. (2021). Influence of organizational innovation support on innovative work behavior of researchers: Chain intermediary effect of creative self-efficacy and knowledge sharing. *Science and Technology Management Research*, *8*, 124–131. https://doi.org/10.3969/j.issn.1000-7695.2021.08.017

Xu, Z., & Suntrayuth, S. (2022). Innovative work behavior in high-tech enterprises: Chain intermediary effect of psychological safety and knowledge sharing. *Frontiers in Psychology*, *13*, 1–11. https://doi.org/10.3389/fpsyg.2022.1017121

Zhang, J., Jehangir, F. N., Yang, L., Tahir, M. A., & Tabasum, S. (2025). Competitive advantage and firm performance: The role of organizational culture, organizational innovation, and knowledge sharing. *Journal of the Knowledge Economy*, *16*(1), 3081–3107. https://doi.org/10.1007/s13132-024-01910-3

Zhu, Y., Liu, J., Lin, S., & Liang, K. (2022). Unlock the potential of regional innovation environment: The promotion of innovative behavior from the career perspective. *Journal of Innovation & Knowledge*, *7*(3), 100206. https://doi.org/10.1016/j.jik.2022.100206

Reporting practices in Romania: Propelling corporate sustainability forward through firm performance

Dorina Nita [1]***, Lia Alexandra Baltador*** [2]***, Codruta Cornelia Dura*** [3]***, Elias Appiah-Kubi*** [4]

[1] *University of Petrosani, Faculty of Economics, Administrative and Social Sciences, Department of Economics, Romania, ORCID: 0000-0002-2639-4971, DorinaNita@upet.ro;*
[2] *University Lucian Blaga of Sibiu, Faculty of Business, Department of Management, Marketing, Business Administration, Romania, ORCID: 0000-0003-3693-4086, lia.baltador@ulbsibiu.ro;*
[3] *University of Petrosani, Faculty of Economics, Administrative and Social Sciences, Department of Economics, Romania, ORCID: 0000-0001-6211-0334, codrutadura@upet.ro;*
[4] *University of Petrosani, Faculty of Economics, Administrative and Social Sciences, Department of Economics, Romania, ORCID:* 0000-0003-4305-2305, *EliasAppiah-Kubi@upet.ro (corresponding author).*

Abstract: *In the context of growing global emphasis on sustainable business practices, understanding the financial drivers of corporate sustainability has become increasingly important. This study examines key financial and market indicators to understand their influence on corporate sustainability disclosure. Focusing on Romanian companies, it investigates how market share, financial ratings, probability of insolvency, and market value shape sustainability outcomes. The dataset was sourced from Azores Sustainability and CSR Services, a well-known organization in Romania that offers insights into the Romania Corporate Sustainability and Transparency Index, as well as from the Risco database. Over an eight-year period from 2016 to 2023, we compiled a total of 264 observations from 33 companies that provided sustainability information. Correlation and regression analysis were used as inferential statistics to achieve the study's objectives. The findings reveal that market share has a positive and significant impact on corporate sustainability disclosure. Additionally, the study shows that financial ratings significantly and positively influence corporate sustainability. The study's results indicate that probability of insolvency has a negative and significant effect on corporate sustainability disclosure. Moreover, market value was found to have a significant positive impact on corporate sustainability disclosure. The findings of the study underscore the relevance of firm performance in influencing Romanian companies' sustainability disclosure following the provisions of Corporate Sustainability Reporting Directive (CSRD) and 2014/95/EU Directive. Ultimately, this study proves essential in the context of the CSRD, which plays a critical role in enhancing corporate transparency regarding sustainability practices. The key contribution of this study is to address the ambiguity in existing research by exploring the reverse relationship, how firm performance influences corporate sustainability outcomes.*

Keywords: *Corporate sustainability disclosure, market share, market value, probability of insolvency, corporate sustainability reporting directive.*

JEL Classification: *M14, M41, L25, Q56.*

APA Style Citation: Nita, D., Baltador, L. A., Dura, C. C., & Appiah-Kubi, E. (2026). Reporting practices in Romania: Propelling corporate sustainability forward through firm performance. *E&M Economics and Management*, *29*(2), 76–91. https://doi.org/10.15240/tul/001/2026-5-002

Early Access Publication Date: *April 8, 2026.*

Introduction

The highly competitive business environment necessitates that firms operate sustainably. However, whether corporate sustainability ultimately benefits or harms firms remains uncertain and unresolved in the literature (Coelho et al., 2023). This debate is ongoing, with no definitive conclusion reached yet, prompting sustainability researchers to further explore the impact of corporate sustainability on firm performance (Park, 2023).

While we acknowledge that numerous studies have been conducted on this topic, the results remain inconclusive and ambiguous (Appiah-Kubi et al., 2024). Some researchers have found that corporate sustainability positively impacts firm performance, while others have found the opposite (Le & Ikram, 2022; Úbeda-García et al., 2021). This is due to differing perspectives: some authors argue that corporate sustainability improves a company's image, which in turn boosts firm performance (Ali et al., 2024; Gupta & Gupta, 2020), while others believe it leads to higher operating costs, ultimately reducing firm performance (Appiah-Kubi et al., 2024). Meanwhile, researchers have neglected to explore how firm performance impacts corporate sustainability disclosure, which could help confirm whether the relationship is bi-directional and reduce the ambiguity in existing research. Besides the above inconclusiveness, sustainability practices among firms in technology-oriented countries like Romania are still minimal (Apostu et al., 2023; Niță & Stoicuța, 2025).

In Romania, few companies make sustainability and occupational health and safety (OHS) disclosures (Iordache et al., 2021; Păun et al., 2020). Notwithstanding, the few companies that engage in sustainability are reluctant to make their sustainability reports available (Grosu et al., 2024). As such, new challenges and perspectives for enhancing non-financial reporting and the disclosure of environmental, social, and governance (ESG) indicators have emerged on the development horizon for Romanian public interest entities, as they implement the provisions of Directive 2014/95/EU and the most recent directive, Corporate Sustainability Reporting Directive (CSRD) into the national regulatory framework (Beleneși et al., 2021). However, the question as to whether the implementation of the Directive 2014/95/EU and CSRD has been impacted by their performance is underexplored, if not unexplored (Ştefănescu et al., 2021). On account of this, it is essential to undertake a study to investigate how corporate sustainability disclosure is influenced by firm performance among Romanian companies (Sahlian et al., 2023). Hence, our study examines the effect of various performance dimensions on Romanian companies' corporate sustainability disclosures. Specifically, the paper achieves four objectives. First, it examines the effect of market share on Romanian companies' corporate sustainability disclosure. Second, it assesses the effect of financial rating on Romanian companies' corporate sustainability disclosure. Third, it examines the impact of probability of insolvency on Romanian companies' corporate sustainability disclosure. Fourth, the effect of market value on Romanian companies' corporate sustainability disclosure is assessed.

By achieving the above objectives, the impacts of various financial performance metrics on Romanian companies' corporate sustainability disclosure are discovered. Sustainability disclosure is particularly relevant in the Romanian context due to the country's ongoing efforts to align with European Union (EU) regulatory standards, including Directive 2014/95/EU on non-financial reporting (Ştefănescu et al., 2021). As a post-transition economy, Romania has historically lagged behind more developed EU members in adopting comprehensive sustainability practices (Păun et al., 2020). However, the mandatory nature of these directives compels Romanian firms, especially publicly listed and large companies, to enhance transparency and report on environmental, social, and governance (ESG) issues. Given that Romania is still in the early stages of implementing structured sustainability reporting frameworks (Niță & Stoicuța, 2025), analyzing disclosure practices offers valuable insights into how firms are responding to external regulatory pressure. Moreover, sustainability disclosure provides a measurable and observable proxy for corporate accountability, signaling mechanisms to stakeholders, and readiness for EU-wide comparability. Understanding how firm-level characteristics influence disclosure behavior thus contributes to both academic discourse and practical policymaking in a region experiencing significant institutional change. The remainder of this paper is structured to feature the literature review, materials and methods, results, and discussion.

1 Theoretical background

The literature on sustainability reporting presents a range of theoretical frameworks that attempt to explain why firms engage in such practices. These frameworks are typically classified into two broad paradigms: social-based theories and economic-based theories (Valentinov & Roth, 2024). Rather than viewing these theories as isolated lenses, this study integrates them into a coherent conceptual scaffold that underscores how corporate sustainability disclosure is shaped by both external societal expectations and internal economic imperatives (see Fig. 1).

Social-based theories, including legitimacy theory, stakeholder theory, and institutional theory, highlight the external drivers of sustainability disclosure. Legitimacy theory posits that firms operate within a “social contract” and must align their actions with societal norms to secure stakeholder approval and maintain access to resources (Ogunode, 2022). This theoretical lens underpins the hypothesis that firms with stronger market visibility (such as greater market share) are more likely to engage in sustainability disclosures to enhance or preserve legitimacy. Stakeholder theory, particularly through the stakeholder-resource-based view (SRBV), explains how firms manage the competing interests of various stakeholder groups. Firms with a broader shareholder base may be more inclined to disclose sustainability information to meet these stakeholders’ expectations and maintain support, providing theoretical justification for including shareholder share as a control variable in the research model. Institutional theory complements these perspectives by focusing on isomorphic pressures: mimetic, coercive, and normative, that drive firms to adopt sustainability practices in order to conform to societal and regulatory expectations. In the Romanian context, where sustainability regulations such as the EU’s CSRD are emerging, institutional

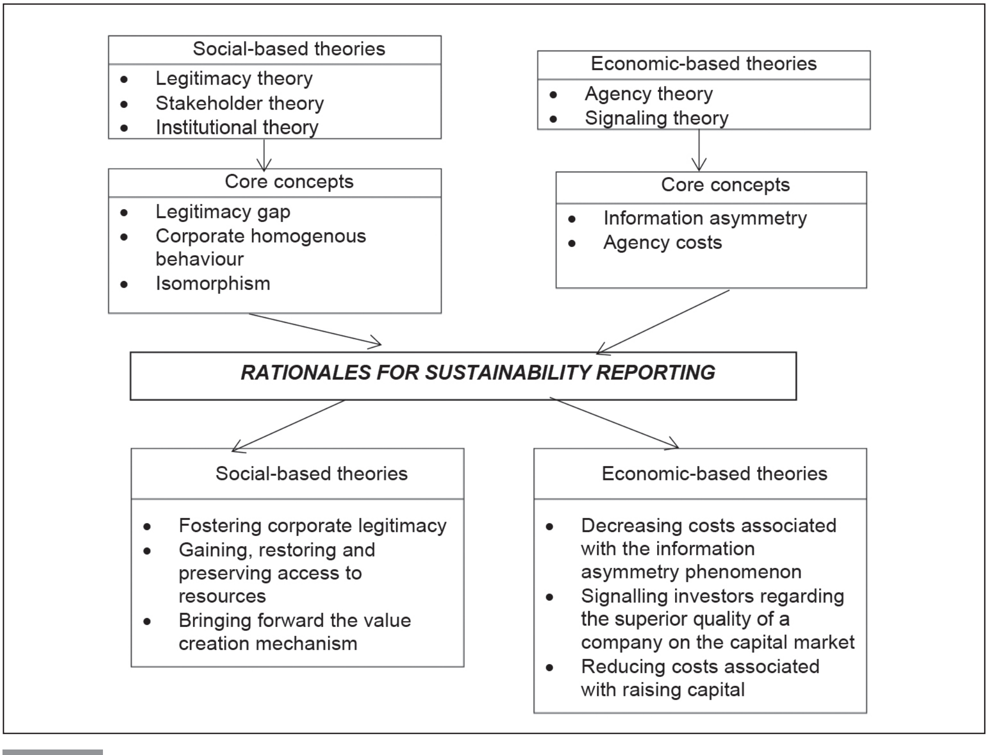

Fig. 1: **Mainstream theories on sustainability reporting**

Source: own

theory supports the rationale for examining how firm-level characteristics like age and eligibility for CST Index evaluation influence sustainability disclosure patterns.

In contrast, economic-based theories such as agency theory and signaling theory explain how sustainability disclosures are used strategically to manage information asymmetries between firms and external stakeholders. Agency theory views such disclosure as a governance mechanism that mitigates opportunistic managerial behavior and aligns managerial actions with shareholder interests. This theory informs the study's expectation that financially stronger firms, as reflected by higher financial ratings, will be more likely to disclose sustainability information to reduce agency costs and increase investor confidence (Vitolla et al., 2020). Signaling theory further suggests that firms engage in sustainability disclosure to convey intangible qualities such as financial stability, responsible management, or long-term viability to stakeholders, particularly investors and regulators. This provides the rationale for examining variables such as market value and probability of insolvency as predictors of disclosure behavior, with the assumption that firms use sustainability reporting to signal strength, resilience, and trustworthiness in a competitive and increasingly transparent business environment (Friske et al., 2023).

Overall, the theoretical framework provides the conceptual foundation for the study's design. Social-based theories justify the inclusion of variables related to institutional legitimacy and stakeholder engagement, while economic-based theories explain the role of firm performance in shaping disclosure behaviors. Together, these perspectives offer a well-rounded and contextually appropriate framework for understanding corporate sustainability disclosure in Romania's transitioning regulatory landscape.

While previous studies have leaned heavily on legitimacy, stakeholder, and institutional perspectives to explore why firms report on sustainability, much of this research has focused on the impact of sustainability on firm performance, typically using CSR disclosures as proxies. A substantial body of work has found a positive association, suggesting that sustainability practices can boost competitiveness, reputation, access to finance, and operational efficiency (Gupta & Gupta, 2020; Ortiz-Martínez et al., 2023). Xu et al. (2020), for instance, analyzed over 2,900 CSR reports from Chinese firms and demonstrated how sustainability disclosures generated value-relevant information, mitigated reputational risk, and enhanced firm value. Similar patterns are evident in empirical studies across emerging and developed markets (Adu-Yeboah et al., 2022; Appiah-Kubi et al., 2024; Ghardallou, 2022).

However, a countervailing stream argues that sustainability efforts may detract from firm performance due to high implementation costs and resource constraints, particularly in economies with limited institutional support (Zhang & Lucey, 2022). The conflicting results point to a methodological and theoretical gap in the literature: most studies adopt a unidirectional causality, exploring how sustainability influences firm outcomes without adequately investigating the reverse relationship: how financial performance indicators such as market share, firm value, or profitability might drive sustainability disclosures.

This study seeks to address this gap by adopting a reverse-causality perspective grounded in signaling and resource-based arguments, thereby repositioning financial performance as a potential antecedent rather than an outcome of sustainability practices. This is especially relevant in the Romanian context, where sustainability reporting is still emerging (Păun et al., 2020), and compliance with EU directives like 2014/95/EU and the CSRD is progressively reshaping corporate disclosure expectations (Beleneși et al., 2021). Despite advancements in Romania's digital infrastructure and regulatory alignment with the EU (Apostu et al., 2023), corporate sustainability reporting remains fragmented and uneven. The country's evolving policy environment, therefore, offers a unique empirical setting to examine how firm-level economic indicators influence sustainability disclosure in a landscape where both regulatory pressures and market mechanisms are still consolidating.

2 Research methodology

2.1 Data and sample

The dataset was sourced from Azores Sustainability and CSR Services, a well-known organization in Romania that offers insights into the Romania Corporate Sustainability and Transparency Index (CST Index), as well as from the Risco database. This index is based on ten sustainability indicators, including sustainability governance, diversity policy, economic impact, climate change and energy, environmental governance, employee responsibility, anti-corruption policies, marketing and awareness initiatives, community investments, supply chain management, and material aspects (management approach disclosures). Romania was deemed suitable for this study due to the limited engagement and reporting of sustainability practices by companies (Beleneși et al., 2021).

This study employs panel data spanning an eight-year period (2016–2023) and focuses on large Romanian companies that publicly disclose sustainability information. At the time of conducting the study, the total number of firms listed in the database was 34. These firms were identified through Azores Sustainability and CSR Services, a reputable organization in Romania known for its evaluation of companies under the CST Index. One company was excluded due to substantial missing data across key variables, resulting in a final sample of 33 firms. These companies operate across diverse sectors, including energy, manufacturing, telecommunications, finance, and consumer goods, offering a representative cross-section of Romania's corporate landscape. Importantly, the dataset exclusively comprises large firms, as sustainability disclosures in Romania remain largely the domain of sizable corporations due to regulatory pressures and stakeholder expectations. The final dataset consists of 264 firm-year observations (33 companies × 8 years), providing adequate variation for robust statistical analysis.

Sustainability disclosure data were primarily sourced from Azores Sustainability and CSR Services, which monitors non-financial corporate reporting in Romania. Financial performance data and other firm-specific variables were obtained from the Risco database, a trusted platform for Romanian financial and credit information. The integration of these data sources enhances both the credibility and comprehensiveness of the dataset. To preserve confidentiality and uphold ethical research standards, the identities of the firms have been anonymized using coded letter identifiers (e.g., Company A, Company B). This approach ensures data integrity while protecting sensitive corporate information.

We acknowledge that the sample comprises only 33 firms; however, these firms represent the entire population of large, publicly listed companies covered by Azores Sustainability and CSR Services for which complete data on sustainability disclosures, financial ratings, and market performance indicators were available during the study period. Although relatively small, the sample is highly representative and contextually significant, reflecting the early and evolving nature of sustainability reporting in Romania. While the limited sample size may constrain the generalizability of the findings, the study offers important insights into how firms in an emerging economy are responding to evolving regulatory requirements, particularly under frameworks such as the EU's CSRD. The findings provide actionable guidance for Romanian companies, policymakers, and other post-transition economies navigating similar regulatory transformations.

2.2 Measurement constructs

The independent variable for this study is firm performance, which captures the financial and market success of a business. According to Mikalef et al. (2023), organizational performance reflects profitability, operational growth, and the capacity to compete effectively in dynamic markets. Firm performance is a central indicator of an organization's health and long-term sustainability. In line with prior studies (e.g., Alsaifi et al., 2020), we operationalized firm performance using two market-based indicators: market share and market value, and two accounting-based indicators: financial rating and probability of insolvency. Market share was measured as a relative percentage of total industry sales, capturing a firm's competitive standing within its sector. Market value was reported in absolute Euro terms and reflects the firm's valuation as perceived by investors. Financial rating indicates the firm's creditworthiness, a proxy for financial health, expressed as a relative score or category. The probability of insolvency was measured in percentage terms and reflects the estimated likelihood of a firm facing financial distress, thus serving as a risk metric. These indicators were sourced from the Romanian Risco database, a reliable repository of firm-level financial and performance data.

The dependent variable is corporate sustainability disclosure, defined as the extent to which firms communicate environmental, social, and governance (ESG) commitments to stakeholders (Stocker et al., 2020). Disclosure serves as a critical transparency mechanism and a response to institutional pressures such as EU regulatory mandates (e.g., the CSRD). We used data from the Romania Corporate Sustainability and Transparency Index, formerly the Romanian CST Index, which is the country's most comprehensive benchmarking tool for sustainability reporting. The index includes both quantitative metrics and qualitative evaluation across ESG dimensions, targeting primarily large firms (500+ employees), though smaller firms may also participate.

We included two control variables: firm age and shareholder share, based on their theorized influence on corporate sustainability strategies and disclosure behavior (Appiah-Kubi, 2025). Firm age reflects organizational maturity, learning capacity, and institutional memory. Older firms may have more established stakeholder relationships, governance structures, and greater pressure to disclose sustainability efforts (Sánchez Pulido et al., 2022). We measured firm age by calculating the natural logarithm of the number of years since incorporation to reduce skewness due to outliers and nonlinearity. Shareholder share represents ownership concentration, capturing the percentage of total shares held by the largest shareholder. A higher concentration may either incentivize or deter sustainability disclosures, depending on whether dominant shareholders are long-term-oriented or focused on short-term returns. Prior studies (e.g., Zaid et al., 2020) suggest that concentrated ownership can influence disclosure practices due to differences in accountability and stakeholder engagement norms. Together, these constructs provide a robust empirical foundation to investigate how firm-level performance and governance dynamics shape sustainability disclosure practices in Romania's evolving institutional context.

2.3 Statistical analysis

To analyze the effect of firm performance indicators on corporate sustainability disclosures, we employed panel data regression using the random effects (RE) model. This choice is grounded in both theoretical and empirical considerations.

First, the structure of our dataset, comprising multiple observations over time for each firm, necessitates an estimation technique that accounts for unobserved heterogeneity across firms. In our case, individual firms may have time-invariant characteristics (such as corporate culture, management philosophy, or reporting history) that could influence disclosure practices. The RE model accounts for this heterogeneity by allowing for a firm-specific error component.

Second, we conducted the Hausman test to determine the appropriate model between fixed effects (FE) and random effects. The null hypothesis of the Hausman test posits that the preferred model is RE, as it assumes no correlation between the firm-specific effects and the explanatory variables. Our test result was statistically insignificant ($p > 0.05$), indicating that the RE model is more efficient and appropriate than the FE model in this context.

Third, the RE model enables the inclusion of time-invariant variables, such as company age, which are important control variables in our analysis. In contrast, fixed-effects models would eliminate these variables due to their inability to estimate coefficients for time-invariant predictors.

Finally, previous studies with similar data structures and research questions, particularly in sustainability and corporate governance research in emerging markets, have widely adopted random-effects estimation (Bahadori et al., 2021). Hence, our methodological choice is aligned with established empirical practices.

The statistics for the variables presented in Tab. 1 indicate that the average CST Index score was 71.01, with a standard deviation of 21.35 from the mean. The median, minimum, and maximum values were 78.00, 22.00, and 99.00, respectively. The CST Index also exhibited a platykurtic distribution, as evidenced by its kurtosis value of less than 3. For the company's age, the average was 23.09, with a standard deviation of 7.58 from the mean, while the median, minimum, and maximum values were 24.00, 3.00, and 33.00, respectively. This variable showed a leptokurtic distribution, with a kurtosis value greater than 3. It could also reflect a mesokurtic distribution due to a skewness value of −0.78, suggesting a potential long left tail with lower values than the sample mean. For the financial rating, the mean was 6.82, with a standard deviation of 1.48 from the mean, while the minimum, median, and maximum values were 2.40, 7.10, and 9.50, respectively. This variable displayed a platykurtic distribution, with a kurtosis value of less than 3 and a skewness of −0.61, indicating the potential for values below the sample mean. In contrast, the market share had a mean of 26.83 and a standard deviation of 28.86 from the mean, with minimum, median and maximum values of 0.004, 0.030, and 0.100, respectively.

Similar to the CST Index, market share exhibited a leptokurtic distribution (with a kurtosis of 3.86), suggesting a peak curve with potentially higher values than the sample mean. The market value averaged 657.33, with a standard deviation of 1,025.50 from the mean. Its minimum, median, and maximum values were 0.06, 14.50, and 100.00, respectively, indicating a leptokurtic distribution (with a kurtosis of 13.93). The probability of insolvency had a mean of 10.71, deviating from the mean by 19.72, with minimum, median, and maximum values of 0.00, 1.00, and 87.00. This also demonstrated a leptokurtic distribution, with a kurtosis of 7.04 and a skewness of 2.22, indicating a peak curve that may include values exceeding the sample mean. Lastly, the shareholder's share had a mean of 1.363 and a standard deviation of 0.73, with its lowest, median, and highest values at 10.40, 15.90, and 22.60, respectively. With a kurtosis value of 3.89, it also exhibited a leptokurtic distribution and a skewness of 1.64, suggesting a peak curve with possible values above the mean. The Jarque-Bera test statistics assesses the differences in skewness and kurtosis of the series compared to a normal distribution, with the null hypothesis (*H0*) stating that the distribution is normal. As shown in Tab. 1, we reject *H0* for the variables since their *p*-values are less than 0.05, indicating that none of the variables follow a normal distribution.

Tab. 1: Variables statistics

	Corporate sustainability and transparency index	Company age	Financial rating	Market share	Market value	Probability of insolvency	Shareholders share
Mean	71.01	23.09	6.82	26.83	657.33	10.71	1.36
Median	78.00	24.00	7.10	14.50	275.50	1.00	1.00
Maximum	99.00	33.00	9.50	100.00	6,500.00	87.00	3.00
Minimum	22.00	3.00	2.40	0.06	0.00	0.00	1.00
Standard deviation	21.35	7.58	1.48	28.86	1,025.50	19.72	0.73
Skewness	−0.70	−0.78	−0.61	1.40	3.00	2.22	1.64
Kurtosis	2.39	3.15	2.72	3.86	13.93	7.04	3.89
Jarque-Bera	12.69	27.18	8.60	45.94	634.34	154.79	127.31
Probability	0.00	0.00	0.01	0.00	0.00	0.00	0.00
Sum	9,160.00	6,096.00	886.20	3,461.06	64,418.67	1,103.00	360.00
Sum of squared deviations	58,356.99	15,125.82	281.18	106,632.60	102,000,000.00	39,653.26	141.09
Observations	129	264	130	129	98	103	264

Source: own

The covariance and correlation results in Tab. 2 reveal that the CST Index has a negative relationship with both company age and probability of insolvency, while exhibiting a stronger positive relationship with market value compared to financial rating and market share, as anticipated. Conversely, the CST Index shows a positive association with financial rating, market share, market value, and shareholders share. However, CST Index shows a negative association with company age, and probability of insolvency. Shareholder's share,

Tab. 2: Covariance and correlation analysis

Covariance and correlation	Corporate sustainability and transparency index	Company age	Financial rating	Market share	Market value	Probability of insolvency	Shareholders share
Corporate sustainability and transparency index	474.76						
	1.00						
Company age	−30.57	60.42					
	−0.18	1.00					
Financial rating	3.51	1.60	1.96				
	0.11	0.15	1.00				
Market share	109.53	33.49	5.07	703.69			
	0.19	0.16	0.14	1.00			
Market value	6,784.83	−125.20	508.11	11,346.74	1,040.93		
	0.31	−0.02	0.36	0.42	1.00		
Probability of insolvency	−1.31	−31.34	−9.73	−105.14	−5,012.86	398.91	
	−0.00	−0.20	−0.35	−0.20	−0.25	1.00	
Shareholders share	1.56	1.33	0.04	14.81	281.42	−3.44	0.56
	0.10	0.23	0.03	0.75	0.37	−0.23	1.00

Source: own

on the other hand, displays a notably strong positive correlation with both market share and market value, as expected. These two variables also show a positive relationship with each other and the financial rating of the companies. Since shareholders own the company and have the right to share in the profits, an increase in market share and market value directly influences shareholder's share. The strongest correlation is observed between shareholder's share and market share. In contrast, shareholder's share has a negative relationship with the probability of insolvency, which aligns with expectations.

The stationarity of the two ratio measures was tested using the Levin, Lin and Chu t^* test (assuming a common unit root process), along with the Im, Pesaran and Shin W-stat, ADF – Fisher chi-square, and PP – Fisher chi-square tests (assuming individual unit root processes). As shown in Tab. 3, the Levin, Lin and Chu t^* and ADF – Fisher chi-square statistics are significant. Therefore, we reject the null hypothesis and conclude that the CST Index has no unit root at the level difference, indicating that it is stationary. Tab. 3 shows the results of the panel unit root test.

2.4 Panel regression analysis

In order to determine the appropriate panel data regression model, pooled ordinary least squares (POLS) was first run in E-Views. The results are presented in Tab. 4.

Given that POLS does not account for unobserved heterogeneity across individual entities, the fixed effect model (FEM) was subsequently estimated. The results are included in Tab. 5.

Thereafter, the random effects model (REM) was estimated. To assess potential endogeneity, the Hausman test was conducted. The test failed to reject the null hypothesis, indicating no significant endogeneity concerns and validating the suitability of the random effects model for the analysis. The Haussmann test hypothesis is as follows:

H_0: *Random effects model is appropriate* (where p-value > 0.05).

Tab. 3: Panel unit root test – CST

Method	Statistic	Probability**	Cross-sections	Observations
Levin, Lin and Chu *t**	−10.68	0.00	2.00	12.00
Breitung *t*-stat	1.35	0.91	2.00	10.00
Im, Pesaran and Shin W-stat	−0.94	0.17	2.00	12.00
ADF – Fisher chi-square	11.17	0.02	2.00	12.00
PP – Fisher chi-square	6.46	0.17	2.00	14.00

Note: ** Probabilities for Fisher tests are computed using an asymptotic chi-square distribution; all other tests assume asymptotic normality. Null: unit root for Levin, Lin and Chu *t** (assumes common unit root process); null: unit root for Im, Pesaran and Shin W-stat (assumes individual unit root process). * indicates the Levin, Lin & Chu test statistic for unit root in panel data.

Source: own

Tab. 4: Pooled Ordinary Least Squares – CST

Variable	Coefficient	Standard error	*t*-statistic	Probability
Constant	68.46	10.07	6.79	0.00
Market share	0.17	0.06	3.16	0.00
Financial rating	4.36	0.94	4.66	0.00
Probability of insolvency	−0.02	0.01	−11.52	0.00
Shareholders share	0.14	0.17	0.85	0.40
Company age	−0.02	0.17	−0.11	0.91
Market value	0.01	0.01	5.94	0.00
R-squared	0.20			
Adjusted *R*-squared	0.18			
Std. error of regression	19.72			
Sum squared residual	99,902.31			
Log likelihood	−1,158.15			
F-statistic	10.74			
Prob(*F*-statistic)	0.00			
Mean dependent variable	66.09			
Std. dev. dependent variable	21.80			
Akaike info criterion	8.83			
Schwarz criterion	8.92			
Hannan-Quinn criterion	8.87			
Durbin-Watson statistic	0.11			

Source: own

Tab. 5: Fixed effect model – CST

Variable	Coefficient	Standard error	*t*-statistic	Probability
Constant	23.64	5.52	4.29	0.00
Market share	0.18	0.06	3.21	0.00
Financial rating	4.19	0.95	4.42	0.00
Probability of insolvency	−0.02	0.01	−17.92	0.00
Shareholders share	0.13	0.07	1.84	0.07
Company age	−0.12	0.18	−0.68	0.50
Market value	0.01	0.01	3.83	0.00
R-squared	0.21			
Adjusted *R*-squared	0.17			
Std. error of regression	19.84			
Sum squared residual	98,419.56			
Log likelihood	−1,156.18			
F-statistic	5.19			
Prob(*F*-statistic)	0.00			
Mean dependent variable	66.09			
Std. dev. dependent variable	21.80			
Akaike info criterion	8.86			
Schwarz criterion	9.05			
Hannan-Quinn criterion	8.94			
Durbin-Watson statistic	0.10			

Source: own

H_a: *Fixed-effect model is appropriate* (where *p*-value < 0.05).

Also, based on the study's objectives, the following model was specified:

$$\begin{aligned} &CST\ Index_{it} = \beta_0 + \beta_1\ market\ value_{it} + \\ &+ \beta_2\ financial\ rating_{it} + \beta_3\ market\ share_{it} + \\ &+ \beta_4\ Probability\ of\ Insolvency_{it} + \\ &+ \beta_5\ Company\ Age_{it} + \\ &+ \beta_6\ Shareholder's\ Share_{it} + \epsilon_{it} \end{aligned} \tag{1}$$

From Tab. 6, the *p*-value is greater than 0.05, indicating no statistical significance, thereby accepting the null hypothesis. Therefore, the random effects model is appropriate for this panel regression for CST.

To ensure the robustness of the results, additional tests were conducted to address potential multicollinearity. A correlation analysis was performed to verify that the independent variables were not highly correlated. As presented in Tab. 2, all correlation coefficients were below the threshold of 0.8, with the highest being 0.749. This confirms the absence of multicollinearity among the variables (Senaviratna & Cooray, 2019).

Moreover, to assess whether there exists cross-sectional dependence among the panel units, three widely used tests were employed: the Breusch-Pagan LM test, the Pesaran scaled LM test, and the Pesaran CD test**.** These tests are particularly relevant in panel data analysis as the presence of cross-sectional dependence can bias the results and lead to inefficient estimators. The results of the cross-sectional dependency tests, as reported in Tab. 7, show that all three tests yielded *p*-values greater

Tab. 6: Hausman test – CST

Test summary		Chi-square statistic	Chi-square degree of freedom	Probability
Cross-section random		1.88	4.00	0.60
Variable	**Fixed**	**Random**	**Var (diff.)**	**Probability**
Financial rating	0.38	0.43	0.03	0.77
Probability of insolvency	0.01	0.02	0.00	0.82
Market value	0.01	0.01	0.00	0.20

Source: own

Tab. 7: Cross-sectional dependence tests

Test	Statistic	Degree of freedom	Probability
Breusch-Pagan LM	948.31	528.00	0.22
Pesaran scaled LM	11.92		0.16
Pesaran CD	9.00		0.18

Source: own

than 0.05. Since none of the *p*-values are statistically significant at the 5% level, we fail to reject the null hypothesis of cross-sectional independence. This indicates that there is no significant cross-sectional dependence among the panel units in the dataset. Therefore, the panel data analysis can proceed under the assumption of cross-sectional independence.

3 Results and discussion

3.1 Results

The random effects model analysis (Tab. 8) demonstrates that key firm performance indicators significantly influence corporate sustainability disclosures (CSD). Specifically, market share shows a positive and statistically significant effect ($\beta = 0.211$; $p < 0.01$), indicating that a 100% increase in market share is associated with a 21.1% increase in CSD, all else being equal. Similarly, financial rating emerges as a strong positive predictor ($\beta = 0.380$; $p < 0.01$), suggesting that improved creditworthiness is linked to greater sustainability disclosure. Conversely, the probability of insolvency has a negative and significant effect on CSD ($\beta = -0.013$; $p < 0.01$), implying that firms facing higher insolvency risks are less likely to engage in sustainability reporting. Market value also contributes positively ($\beta = 0.081$; $p < 0.01$), though with a smaller magnitude, reinforcing the notion that firms with stronger market capitalization are more inclined to disclose sustainability-related information. Regarding the control variables, both shareholders' share ($\beta = 0.060$) and company age ($\beta = 0.075$) exhibit positive but statistically insignificant effects ($p > 0.05$), suggesting limited influence on disclosure practices within the sample. The adjusted R-squared value of 95.99% indicates that the model explains a substantial proportion of the variance in corporate sustainability disclosures, underscoring the robustness of the explanatory variables.

3.2 Discussion

Motivated by the limited number of Romanian companies that disclose sustainability information and the often poor quality of such disclosures (Păun et al., 2020), this study examined how firm performance influences

Tab. 8: Random effect model – CST

Variable	Coefficient	Standard error	*t*-statistic	Probability
Constant	66.37	4.72	14.07	0.00
Market share	0.21	0.04	4.85	0.00
Financial rating	0.38	0.09	4.05	0.00
Probability of insolvency	−0.01	0.01	−4.91	0.00
Shareholders share	0.06	0.07	0.91	0.36
Company age	0.08	0.13	0.59	0.56
Market value	0.08	0.02	3.85	0.00
***R*-squared**	0.97			
Adjusted *R*-squared	0.96			
Std. error of regression	4.39			
Sum squared residual	1,231.54			
Log likelihood	−263.08			
***F*-statistic**	71.33			
Probability (*F*-statistic)	0.00			
Mean dependent variable	69.88			
Std. dev. dependent variable	21.90			
Akaike info criterion	6.06			
Schwarz criterion	6.96			
Hannan-Quinn criterion	6.43			
Durbin-Watson statistic	1.74			

Source: own

sustainability disclosure under the framework of Directive 2014/95/EU and the CSRD. The results reveal that firm performance is a significant determinant of sustainability disclosure across several dimensions. Specifically, market share and market value exhibited a significant positive effect on sustainability disclosure, underscoring the role of competitive positioning and external valuation in motivating firms to engage in sustainability reporting. Additionally, financial rating was found to positively influence disclosure, while the probability of insolvency had a negative effect. These findings highlight that financially healthier firms are more likely to commit to transparent sustainability practices. These results are theoretically supported by signalling theory, suggesting that firms use sustainability disclosure as a strategic tool to communicate financial strength and reliability to stakeholders. Empirically, the study advances the literature by reinforcing the bi-directional relationship between sustainability disclosure and firm performance, where stronger financial performance promotes disclosure, and disclosure, in turn, enhances stakeholder perception and firm value.

Our findings align with prior studies across different contexts. For instance, Lakatos et al. (2021) found that customer perceptions of sustainability influence loyalty in Romania. Similarly, studies in Ghana (Adu-Yeboah et al., 2022; Appiah-Kubi et al., 2024) and Saudi Arabia (Ghardallou, 2022) affirmed the positive relationship between sustainability practices and firm performance. Gupta and Gupta (2020) also confirmed this relationship from the environmental sustainability lens. In sum, this study contributes to the literature by empirically demonstrating that firm performance significantly drives sustainability disclosure in Romania. It highlights the strategic relevance of sustainability reporting as both a reflection of financial strength and a mechanism for enhancing organizational legitimacy, especially in post-transition economies adapting to EU sustainability mandates.

Conclusions

The primary aim of this study was to investigate whether the corporate sustainability disclosure of Romanian companies depends on firm performance, in an effort to determine whether the implementation of 2014/95/EU and CSRD directives has been influenced by their performance. Specifically, the study explores the impact of market share on the corporate sustainability disclosure of Romanian companies. Secondly, it evaluates the influence of financial ratings on their sustainability disclosure. Thirdly, it investigates how the probability of insolvency affects corporate sustainability disclosure. Lastly, the study assesses the effect of market value on the sustainability disclosure of Romanian companies. The main contribution of this study is to reduce the ambiguity in existing research by exploring how firm performance impacts corporate sustainability disclosure.

The findings reveal that market share has a positive and significant impact on corporate sustainability disclosure, underscoring the role of competitive advantage in boosting sustainability outcomes. Additionally, the study shows that financial ratings significantly and positively influence corporate sustainability, highlighting the need to strengthen corporate solvency. This is further supported by the third finding, which indicates that a lower probability of insolvency has a negative and significant effect on sustainability disclosure, suggesting that companies with lower bankruptcy risks engage more in sustainability initiatives. Moreover, market value was found to have a significant positive impact on corporate sustainability disclosure, emphasizing the importance of enhancing a company's market worth.

Given that market share has a positive and significant impact on corporate sustainability disclosure, the board of directors should focus on meeting customer demands to achieve a competitive edge and expand market share. For example, they should provide high-quality products at affordable prices and ensure that their products or services are easily accessible to customers. Furthermore, given the positive and significant impact of financial ratings on corporate sustainability disclosure, managers should aim to enhance both short-term and long-term liquidity. They are encouraged to efficiently manage existing resources to ensure sufficient cash flow to support business operations. Moreover, considering the significant negative effect of insolvency risk on corporate sustainability disclosure, managers should adhere to corporate governance principles to strengthen the firms' long-term viability. In light of the significant positive impact of market value on corporate sustainability disclosure, managers should aim to enhance their firms' value. This can be accomplished by broadening their operations through product development, market expansion, market penetration and both concentric and conglomerate diversification. Additionally, managers are encouraged to engage in regular market research to support benchmarking efforts that could enhance their firms' market value. Furthermore, forming strategic alliances is recommended as a strategy for improving the market value of their companies.

Alongside the managerial implications mentioned, this paper also offers policy recommendations. The Romanian government is urged to establish measures that support the financial performance of local companies. For example, the government could offer funding assistance to businesses dedicated to enhancing their corporate sustainability disclosure. Additionally, it is suggested that the European Union provide financial aid to companies facing liquidity challenges to assist in meeting the requirements of CSRD Directive.

Beyond the Romanian context, the study draws attention to the universal challenges faced by companies in emerging and post-transition economies, particularly regarding their capacity to comply with evolving sustainability reporting frameworks. By empirically examining how firm-level performance indicators such as market share, financial ratings, and solvency risk influence sustainability disclosure among Romanian companies, the study provides contextualized insights that are highly transferable to other countries facing similar structural and institutional conditions. These include nations in Eastern and Central Europe, the Balkans, and other post-socialist or developing EU member states where sustainability reporting is still evolving. The findings emphasize the importance of strengthening firm performance, not only as a driver of economic competitiveness but also as a signal to stakeholders of a firm's sustainability orientation. The results highlight the need to design tailored regulatory interventions and support mechanisms that align firm capabilities with disclosure requirements. This dual relevance positions the study as a meaningful contribution to the broader discourse on advancing corporate sustainability in contexts marked by transition, limited capacity, and regulatory shifts.

While this study makes important contributions to the literature, it has limitations that highlight potential areas for future research. The use of panel data resulted in some missing information. Given that many Romanian companies do not report their sustainability disclosure, which complicates access to their secondary data, future studies should focus on gathering primary data. Utilizing both questionnaires and interviews would yield deeper insights into their sustainability and firm performance. Consequently, future research should address common method bias to explore the impact of firm performance on corporate sustainability disclosure. Although this study analyzed data from 2016 to 2023, future research should extend the time frame to enhance the number of observations and improve the generalizability of the findings. Furthermore, while this study concentrated on large companies in Romania, future research should also include Small and Medium Enterprises as they play a crucial role in contributing to the Sustainable Development Goals given their prevalence in almost every country. Also, future studies might consider companies in other countries in an attempt to generalize the findings of our study.

Acknowledgments: *The authors duly acknowledge the assistance provided by Azores Sustainability and CSR Services for providing data to support the study.*

References

Adu-Yeboah, S. S., Jiang, Y., Frempong, M. F., Hossin, M. A., & Amoako, R. (2022). Corporate sustainability and firm performance in small and medium enterprises in Ghana: Mediating role of green innovation. *Journal of Psychology in Africa*, *32*(4), 311–318. https://doi.org/10.1080/14330237.2022.2066341

Ali, A., Jiang, X., & Ali, A. (2023). Enhancing corporate sustainable development: Organizational learning, social ties, and environmental strategies. *Business Strategy and the Environment*, *32*(4), 1232–1247. https://doi.org/10.1002/bse.3184

Alsaifi, K., Elnahass, M., & Salama, A. (2020). Carbon disclosure and financial performance: UK environmental policy. *Business Strategy and the Environment*, *29*(2), 711–726. https://doi.org/10.1002/bse.2426

Apostu, S. A., Nichita, E. M., Manea, C. L., Irimescu, A. M., & Vulpoi, M. (2023). Exploring the influence of innovation and technology on climate change. *Energies*, *16*(17), 6408. https://doi.org/10.3390/en16176408

Appiah-Kubi, E. (2025). Stakeholder pressure and SMEs sustainability information disclosures: Why perceived benefits and sustainability knowledge matter. *Journal of Accounting & Organizational Change*, *21*(3), 407–423. https://doi.org/10.1108/JAOC-11-2023-0191

Appiah-Kubi, E., Boateng, R. N., Dogbe, C. S. K., & Kumah, S. P. (2024). Organisational sustainability and SMEs performance: The role of control environment. *Journal of Cleaner Production*, *452*, 142026. https://doi.org/10.1016/j.jclepro.2024.142026

Bahadori, N., Kaymak, T., & Seraj, M. (2021). Environmental, social, and governance factors in emerging markets: The impact on firm performance. *Business Strategy & Development*, *4*(4), 411–422. https://doi.org/10.1002/bsd2.167

Beleneși, M., Bogdan, V., & Popa, D. N. (2021). Disclosure dynamics and non-financial reporting analysis. The case of Romanian listed companies. *Sustainability*, *13*(9), 4732. https://doi.org/10.3390/su13094732

Coelho, R., Jayantilal, S., & Ferreira, J. J. (2023). The impact of social responsibility on corporate financial performance: A systematic literature review. *Corporate Social Responsibility and Environmental Management*, *30*(4), 1535–1560. https://doi.org/10.1002/csr.2446

Directive 2014/95/EU of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups, OJ L 330, 15.11.2014, p. 1–9.

Friske, W., Hoelscher, S. A., & Nikolov, A. N. (2023). The impact of voluntary sustainability reporting on firm value: Insights from signaling theory. *Journal of the Academy of Marketing Science*, *51*(2), 372–392. https://doi.org/10.1007/s11747-022-00879-2

Ghardallou, W. (2022). Corporate sustainability and firm performance: The moderating role of CEO education and tenure. *Sustainability*, *14*(6), 3513. https://doi.org/10.3390/su14063513

Grosu, V., Ionescu-Feleagă, L., Macovei, A.-G., Ciubotariu, M.-S., Hlaciuc, E., Socoliuc, M., & Petrescu, C. (2024). The impact of IFRS 9 adoption on the financial performance and sustainability of Romanian credit institutions. *Eastern European Economics*, *63*(6), 993–1014. https://doi.org/10.1080/00128775.2024.2368031

Gupta, A. K., & Gupta, N. (2020). Effect of corporate environmental sustainability on dimensions of firm performance – Towards sustainable development: Evidence from India. *Journal of Cleaner Production*, *253*, 119948. https://doi.org/10.1016/j.jclepro.2019.119948

Iordache, A. M. M., Dura, C. C., Coculescu, C., Isac, C., & Preda, A. (2021). Using neural networks in order to analyze telework adaptability across the European Union countries: A case study of the most relevant scenarios to occur in Romania. *International Journal of Environmental Research and Public Health*, *18*(20), 10586. https://doi.org/10.3390/ijerph182010586

Lakatos, E.-S., Nan, L.-M., Bacali, L., Ciobanu, G., Ciobanu, A.-M., & Cioca, L.-I. (2021). Consumer satisfaction towards green products: Empirical insights from Romania. *Sustainability*, *13*(19), 10982. https://doi.org/10.3390/su131910982

Le, T. T., & Ikram, M. (2022). Do sustainability innovation and firm competitiveness help improve firm performance? Evidence from the SME sector in Vietnam. *Sustainable Production and Consumption*, *29*, 588–599. https://doi.org/10.1016/j.spc.2021.11.008

Mikalef, P., Islam, N., Parida, V., Singh, H., & Altwaijry, N. (2023). Artificial intelligence (AI) competencies for organizational performance: A B2B marketing capabilities perspective. *Journal of Business Research*, *164*, 113998. https://doi.org/10.1016/j.jbusres.2023.113998

Niță, D., & Stoicuța, N. (2025). The impact of energy consumption on economic growth: Application of CES function for Romania. *Journal of Business Economics and Management*, *26*(2), 444–464. https://doi.org/10.3846/jbem.2025.23697

Ogunode, O. A. (2022). Legitimacy theory and environmental accounting reporting and practice: A review. *South Asian Journal of Social Studies and Economics*, *13*(1), 17–28. https://doi.org/10.9734/SAJSSE/2022/v13i130345

Ortiz-Martínez, E., Marín-Hernández, S., & Santos-Jaén, J.-M. (2023). Sustainability, corporate social responsibility, non-financial reporting and company performance: Relationships and mediating effects in Spanish small and medium sized enterprises. *Sustainable Production and Consumption*, *35*, 349–364. https://doi.org/10.1016/j.spc.2022.11.015

Park, S.-B. (2023). Bringing strategy back in: Corporate sustainability and firm performance. *Journal of Cleaner Production*, *388*, 136012. https://doi.org/10.1016/j.jclepro.2023.136012

Păun, A. P., Dura, C. C., Mihăilescu, S., Moraru, R. I., & Isac, C. A. (2020). OHS disclosures within non-financial reports: The Romanian case. *Sustainability*, *12*(5), 1963. https://doi.org/10.3390/su12051963

Sahlian, D. N., Popa, A. F., Banţa, V. C., Răpan, C. M., & Chiriac, S. C. V. (2023). The value relevance of financial performance and ESG scores during tumultuous economic periods. *Eastern European Economics*, *62*(4), 528–546. https://doi.org/10.1080/00128775.2023.2225500

Sánchez Pulido, L., Moreno Gené, J., & Gallizo Larraz, J. L. (2022). Internationalization of family firms: The effect of CEO attributes. *Journal of Management and Governance*, *26*(4), 1123–1154. https://doi.org/10.1007/s10997-021-09597-3

Senaviratna, N. A. M. R., & Cooray, T. M. J. A. (2019). Diagnosing multicollinearity of logistic regression model. *Asian Journal of Probability and Statistics*, *5*(2), 1–9. https://doi.org/10.9734/AJPAS/2019/v5i230132

Ştefănescu, C. A., Tiron-Tudor, A., & Moise, E. M. (2021). EU non-financial reporting research – Insights, gaps, patterns and future agenda. *Journal of Business Economics and Management*, *22*(1), 257–276. https://doi.org/10.3846/jbem.2020.13479

Stocker, F., de Arruda, M. P., de Mascena, K. M. C., & Boaventura, J. M. G. (2020). Stakeholder engagement in sustainability reporting: A classification model. *Corporate Social Responsibility and Environmental Management*, *27*(5), 2071–2080. https://doi.org/10.1002/csr.1947

Úbeda-García, M., Claver-Cortés, E., Marco-Lajara, B., & Zaragoza-Sáez, P. (2021). Corporate social responsibility and firm performance in the hotel industry. The mediating role of green human resource management and environmental outcomes. *Journal of Business Research*, *123*, 57–69. https://doi.org/10.1016/j.jbusres.2020.09.055

Valentinov, V., & Roth, S. (2024). Stakeholder theory: Exploring systems-theoretic and process-philosophic connections. *Systems Research and Behavioral Science*, *41*(2), 301–315. https://doi.org/10.1002/sres.2970

Vitolla, F., Raimo, N., & Rubino, M. (2020). Board characteristics and integrated reporting quality: An agency theory perspective. *Corporate Social Responsibility and Environmental Management*, *27*(2), 1152–1163. https://doi.org/10.1002/csr.1879

Xu, S., Chen, X., Li, A., & Xia, X. (2020). Disclosure for whom? Government involvement, CSR disclosure and firm value. *Emerging Markets Review*, *44*, 100717. https://doi.org/10.1016/j.ememar.2020.100717

Zaid, M. A. A., Abuhijleh, S. T. F., & Pucheta-Martínez, M. C. (2020). Ownership structure, stakeholder engagement, and corporate social responsibility policies: The moderating effect of board independence. *Corporate Social Responsibility and Environmental Management*, *27*(3), 1344–1360. https://doi.org/10.1002/csr.1888

Zhang, D., & Lucey, B. M. (2022). Sustainable behaviors and firm performance: The role of financial constraints' alleviation. *Economic Analysis and Policy*, *74*, 220–233. https://doi.org/10.1016/j.eap.2022.02.003

Profitability index in the accommodation services sector in Slovakia

Jana Spirkova[1], *Miroslava Vinczeova*[2]

[1] Matej Bel University in Banská Bystrica, Faculty of Economics, Department of Quantitative Methods and Information Systems, Slovakia, ORCID: 0000-0001-5864-9353, jana.spirkova@umb.sk (corresponding author);
[2] Matej Bel University in Banská Bystrica, Faculty of Economics, Department of Corporate Economics and Management, Slovakia, ORCID: 0000-0001-5380-0907, miroslava.vinczeova@umb.sk.

Abstract: *Considering evolving market dynamics, high fixed costs and vulnerability to external shocks, such as the impact of the ongoing global pandemic, it has become increasingly important to assess profitability in the accommodation sector. This study proposes a composite profitability index for accommodation services in Slovakia, which integrates traditional financial indicators such as return on equity (ROE), return on assets (ROA) and return on sales (ROS), as well as broader indicators such as the value-added-to-sales ratio (VAS) and EBITDA margin. Each indicator captures a distinct aspect of financial health, from operational efficiency to capital utilisation and value creation. To ensure methodological robustness, the study employs three expert-based weighting techniques: the exact Saaty method, the approximate Saaty method and the best-worst method. A panel of 14 experts in economics, finance and hospitality management participated in a Delphi-based survey to evaluate the relative importance and optimal values of each indicator. All methods ranked ROA as the most critical metric, followed by ROS and ROE, meanwhile, EBITDA margin and VAS were considered less significant, yet still essential, for a comprehensive assessment. Using sectoral data and applying normalised scoring, the index captures profitability trends from 2020 to 2023. The aim of the study is to develop a composite profitability index for the accommodation services sector in Slovakia, based on the identified limitations of traditional indicators. The results show a marked post-pandemic recovery, with profitability index scores rising steadily across all weighting methods. Notably, the best-worst method yielded consistently higher index values, indicating its sensitivity to prioritisation extremes. This composite index provides accommodation businesses and policymakers with a practical tool for benchmarking, strategic planning and improving long-term financial sustainability in a volatile and competitive environment, offering a multidimensional framework for evaluating financial performance.*

Keywords: *Indicators, weights, exact Saaty method, approximate Saaty method, best-worst method.*

JEL Classification: *B41, C10, Z30, Z32.*

APA Style Citation: Spirkova, J., & Vinczeova, M. (2026). Profitability index in the accommodation services sector in Slovakia. *E&M Economics and Management*, *29*(2), 92–106. https://doi.org/10.15240/tul/001/2026-2-007

Introduction

The accommodation sector plays a crucial role in each economy, particularly in countries with strong tourism potential such as Slovakia. It contributes not only to GDP generation, but also significantly supports regional development, employment and the overall attractiveness of the country. However, in recent years the sector has faced unprecedented challenges brought about by turbulent developments in external environment. The COVID-19 pandemic led to a sharp decline in tourism and occupancy rates, resulting in severe financial distress for many entrepreneurs in the sector. Although recovery is underway, new issues continue to emerge, including changing customer behaviour, increasing operational costs, and labour shortages. In Slovakia, the industry has also been impacted by legislative changes, such as shifts in the VAT rate for selected tourism services and proposed changes in taxation and wage policy, including the debated introduction of a controversial transaction tax. Additionally, rising energy prices and inflation have put further pressure on the cost structures and profitability of accommodation businesses. These dynamics highlight the need to reassess the financial sustainability of enterprises operating in this sector characterized by their high fixed costs and variable demand. Under these conditions, a nuanced understanding of how profitability can be meaningfully assessed is essential for both practical decision-making and informed policy responses aimed at strengthening the resilience of the accommodation industry in the long term.

Profitability represents a key dimension in assessing a company's financial performance. In the hospitality sector, particularly within accommodation services, profitability analysis involves distinct challenges due to high fixed costs and demand volatility (Alrawabdeh, 2021). Unlike liquidity and indebtedness, which reflect financial stability and capital structure, profitability focuses on the performance and efficiency of business operations. It expresses the firm's ability to effectively utilize available capital, assets, or revenues to generate profit (Fadhillah et al., 2024). The importance of profitability metrics is highlighted by their capacity to guide strategic decision-making, attract investment, and facilitate performance benchmarking against competitors (Santos et al., 2022). Moreover, regular profitability analysis provides a more precise understanding of how qualitative outcomes evolve in response to the transformation of business processes, reflecting the efficiency of managing inputs and outputs (Lesáková et al., 2019). While profitability indicators offer valuable insights into the financial performance of enterprises, their interpretation must account for the broader external environment (Costa & Costa, 2024). The accommodation services sector is particularly exposed to such influences, as demonstrated by the COVID-19 pandemic, when travel restrictions and lockdowns led to substantial revenue losses (Blahušiaková, 2021; Vojteková & Klieštik, 2024). These circumstances highlight the sector's vulnerability to environmental disruptions, which markedly affect profitability even in otherwise well-performing firms (Šebová et al., 2024).

Traditional profitability indicators, such as return on assets (ROA), return on equity (ROE), and return on sales (ROS), provide only a partial view of a firm's financial health, with their interpretation often influenced by capital structure, accounting policies, and the tax environment (Altman, 1968; Beaver, 1966; Ohlson, 1980). Existing studies predominantly focus on individual indicators, resulting in a fragmented perspective on financial performance. Contemporary authors (Jović & Tomašević, 2021; Santos et al., 2022) therefore emphasize the need to extend performance assessment to include indicators capturing broader aspects of performance, such as the EBITDA margin or the share of value added in revenues. Despite these efforts, there remains a lack of an integrated tool that consolidates these indicators into a unified framework, enabling a multidimensional evaluation of profitability. The study addresses this gap by introducing a composite profitability index specifically designed for the accommodation services sector. The index integrates traditional financial indicators (ROE, ROA, ROS) with extended metrics, such as the share of value added in revenues and the EBITDA margin, thereby providing a more comprehensive assessment of firm financial performance.

To ensure methodological rigour, the study employs three expert-based weighting techniques: the exact Saaty method, the approximate Saaty method and the best-worst method. Drawing on median industry data over the period 2020–2023, the findings reveal a consistent post-pandemic recovery trend and demonstrate the usefulness of a multidimensional framework for assessing profitability. The proposed index has valuable implications for hospitality practitioners, policymakers and tourism stakeholders who are seeking to improve strategic planning, benchmarking and financial decision-making in an increasingly complex operating environment.

The aim of this study is to develop a composite profitability index for the accommodation services sector in Slovakia, based on the identified limitations of traditional indicators. The proposed index integrates multiple financial metrics, enabling a more comprehensive assessment of financial performance. Based on this aim, the research question is formulated as follows: Which profitability indicators are considered most significant by experts, what weights should they carry in constructing a composite index, and how does the resulting index vary depending on the weighting method applied?

The paper is organized as follows. Chapter 1 – Role of profitability in business performance: Indicators and sector-specific considerations, gives information about all indicators examined. Chapter 2 – Methodology and expert-based weighting results provides methodology and information on determining weights using exact Saaty method, approximate Saaty method and so-called best-worst method. Chapter 3 – Results and discussion: Overall profitability index presents the results of our research and show the growth trend of the profitability index in accommodation sector in Slovakia. Finally, Conclusions suggest further avenues for investigating the profitability of not only accommodation services, but also new methods of determining weights in multi-criteria decision-making.

1 Role of profitability in business performance: Indicators and sector-specific considerations

The conceptual roots of modern profitability metrics reach back to the early 20th century. Brown's DuPont framework, developed around 1912, introduced return on investment (ROI) as a key measure of business performance – functionally equivalent to what is now commonly referred to as return on assets (ROA). By decomposing ROI into profitability and efficiency components, Brown provided the analytical foundation upon which later ratio-based performance frameworks were built (Flesher & Previts, 2013). Over subsequent decades, numerous scholars (i.e., Altman, 1968; Beaver, 1966; Ohlson, 1980) further advanced the empirical relevance of profitability indicators. Conventional approaches to profitability assessment emphasize the calculation of several key financial ratios, with return on equity (ROE), return on assets (ROA), and return on sales (ROS) being most frequently employed indicators. ROE assesses profitability relative to owners' equity, making it a critical metric for attracting and retaining investors. Conversely, ROA is a robust indicator of a company's efficiency in utilizing its assets to generate earnings (Aqil et al., 2019). ROA is influenced by the effective utilization of assets, with performance tied closely to occupancy rates and revenue per available room (Mundi et al., 2023). It can be calculated using various profit measures depending on the analytical objective and context. The most common formula uses net income in the numerator providing a comprehensive view of profitability after interest and taxes. Alternatively, EBIT (earnings before interest and taxes) is often used to assess operational efficiency, as it excludes the effects of financing and taxation. Some financial databases, such as cribis.sk (2025), apply EBT (earnings before taxes) to neutralize differences in tax regimes across companies. In performance analysis and value-based management, NOPAT (net operating profit after taxes) may be preferred for its focus on after-tax operational returns (Bouwens et al., 2019). Selecting the appropriate version of ROA depends on the purpose of the analysis – whether evaluating overall profitability, isolating operating performance, or enabling cross-company comparability. A higher ROA indicating efficient asset utilization in the hospitality sector, where asset-heavy investments are common, is crucial given the capital-intense nature of the industry (Mundi et al., 2023). ROS, also referred to as operating profit margin when based on operating income, captures the efficiency of a firm's core operational processes, while remaining unaffected by various secondary influences, many of which are external and beyond the firm's control. It measures the percentage of revenue that translates into operating profit (Lee, 2024). The hospitality industry relies on ROS to gauge both operational performance and cost-efficiency (Jugović & Maričić, 2024).

Despite certain conceptual limitations compared to more modern approaches, traditional profitability indicators continue to offer several important advantages. They satisfy several fundamental criteria for effective performance measurement, including accuracy, objectivity, and comprehensibility. Their values can typically be influenced by managerial actions and measured over relatively short time periods. Furthermore, they are grounded in internationally recognized accounting and financial reporting standards (Jović & Tomašević, 2021). Their auditability enhances their credibility and ensures their continued relevance in both academic research and practical financial analysis. These ratios maintain a prominent role in firm's performance evaluation, largely owing to their interpretability, wide applicability, and the simplicity of their computation based on readily accessible financial statement data. This facilitates consistency and comparability across firms and industries, while also serving as a foundation for investment decision-making (Palepu et al., 2022). On the other hand, it is important to acknowledge the limitations of these indicators, particularly their high sensitivity to accounting and tax policies, as well as the firm's capital structure, Damodaran (2014). Contemporary research highlights the need to broaden the traditional understanding of profitability by incorporating indicators that capture wider aspects of value creation. Increasingly, additional measures such as the value-added to revenue ratio (VAS) and the EBITDA margin are being employed in profitability assessments. The VAS ratio is important mainly in assessing the productivity and impact of a business on its wider economic environment. It assesses the extent to which a firm internally generates value relative to its sales revenue. It expresses the proportion of sales that remains after deducting the cost of external inputs, thereby reflecting the firm's capacity to generate economic value through its operations. A higher VAS ratio indicates a firm's greater ability to retain value within the production process, through labour, capital, and innovation, rather than transferring it to suppliers. The VAS ratio complements traditional metrics by offering insight into the structural efficiency of the firm, particularly in terms of its value-generating potential (Tangen, 2005). Since value added reflects the income generated by a firm before it is distributed to capital and labour, it serves as a broader measure of operational performance that is less affected by differences in financial structure or tax regimes. In both cross-industry and longitudinal analyses, the VAS ratio can be employed to assess firms' underlying performance, particularly in industries where the cost of materials and intermediate goods represents a substantial share of total revenue. This ratio highlights the hospitality sector's reliance on service quality and guest experience as drivers of value creation.

The EBITDA margin provides a more comprehensive view of a firm's financial health (Setyopurnomo et al., 2025), capturing core operational performance without the distortion caused by tax strategies, depreciation methods, or variations in capital structure. EBITDA serves as a generalizable metric (a "one-size-fits-all" accounting figure) that summarizes a firm's profitability, cash-generating capacity, and its ability to service debt. It is calculated by adding back depreciation and amortization, which are non-cash expenses that do not affect the company's current cash flow (Bouwens et al., 2019). Additionally, some argue that excluding interest and taxes enhances the comparability of the metric across different firms (D'Souza et al., 2010). The EBITDA margin has emerged as a widely recognized proxy for assessing long-term operational efficiency, Maxim (2023). In recent years, it has attracted considerable attention both in academic literature and in practice, owing to its capacity to reflect core profitability unaffected by capital structure, tax regimes, or non-cash accounting items.

As such, it is frequently employed as a standardized benchmark for comparing profitability across firms and industries. A high EBITDA margin signals strong core profitability, critical for covering fixed costs and supporting ongoing investment in property and service improvements.

Subsequently, the formulas for computing the above-mentioned profitability ratios are introduced as follows:

$$ROE = \frac{\text{net income}}{\text{shareholder equity}} \times 100\% \tag{1}$$

$$ROA = \frac{\text{net income}}{\text{total assets}} \times 100\% \tag{2}$$

$$ROS = \frac{\text{operating profit}}{\text{net sales}} \times 100\% \tag{3}$$

$$\begin{matrix}\text{Share of value}\\ \text{added in sales}\end{matrix} = \frac{\text{operating profit}}{\text{net sales}} \times 100\% \tag{4}$$

$$\begin{matrix}\text{EBITDA}\\ \text{margin}\end{matrix} = \frac{\begin{matrix}\text{net income} + \text{taxes} +\\ + \text{interest expense} +\\ + D\&A\end{matrix}}{\text{net sales}} \times 100\% \tag{5}$$

where: *D&A* – depreciation and amortization.

Understanding profitability necessitates a comprehensive financial analytical approach that integrates multiple key performance indicators (Elexa, 2021). Profitability metrics provide crucial insights into operational efficiency, capital utilization, and overall financial health. Employing these indicators enables firms to diagnose performance drivers and areas of strength or weakness, thereby facilitating more informed strategic decision-making and promoting sustainable growth. However, relying solely on individual financial ratios can offer fragmented perspective that may not fully capture the multidimensional nature of firm's performance (Altman, 1968; Beaver, 1966). To address these limitations, the development of a composite index construed using expert-based methodologies offers a more integrated and robust assessment framework in multi-criteria decision making (Ishizaka & Labib, 2011; Saaty, 1977). By aggregating diverse financial metrics into a weighted index, tailored specifically for the hospitality accommodation sector, this approach contributes to a broader understanding of financial performance by combining key indicators into a unified index (Ohlson, 1980). Such composite indices are invaluable for benchmarking and comparative analysis, enabling stakeholders to identify best practices and emerging risks within the industry. This is particularly critical in hospitality, where market volatility, seasonality, and heterogeneous business models pose challenges for straightforward financial evaluation (Enz, 2009). Incorporating expert insights to determine the relative importance and optimal values of key profitability metrics enhances the index's relevance and practical utility of these metrics within the sector's specific context.

The profitability index is a comprehensive measure combining various profitability indicators to provide a more holistic view of a firm's financial health. In the hospitality sector, this index can help benchmarking performance against industry standards and identifying areas for improvement.

2 Methodology and expert-based weighting results

To construct a profitability index for the accommodation services sector, we first needed to obtain data on the perceived importance of selected profitability indicators and their values considered to be acceptable ("optimal") from the perspective of long-term business sustainability in this sector. To achieve this, we approached 20 experts. Their selection was based on their demonstrated expertise in economics, finance and financial analysis, accounting, business, and accommodation management, as well as on their practical experience in assessing financial performance or on relevant publications in the field. From the experts contacted, we ultimately formed a panel of 14 experts who expressed willingness to collaborate with us. The expert panel comprised 14 participants: six academics specializing in financial analysis and tourism, four managers of accommodation facilities with over ten years of professional experience, two hospitality consultants, and two professionals from industry associations. Data were collected in two rounds of a modified Delphi method between March and May 2025, using a structured electronic assessment form. Each round lasted approximately two weeks, with anonymized summaries of responses provided to participants between rounds for review. The form consisted of two main sections: (i) pairwise comparison of five indicators using the analytic hierarchy process (AHP) scale (1–9); and (ii) estimation of the optimal values for the indicators. Each section was accompanied by a methodological explanation and an illustrative example. The assessment form was based on the procedures described by Saaty (1977) and Rezaei (2016a) and was validated through pilot testing. The participating experts were asked to respond to two main questions, following a detailed explanation of the research objectives and the response procedure.

Determining the relative importance of profitability indicators

Respondents were asked to rank five selected profitability indicators based on their subjective assessment of the indicators' relevance for evaluating the sustainability and performance of accommodation service providers. The indicators were: I1 – return on equity (ROE); I2 – return on assets (ROA); I3 – return on sales (ROS); I4 – value added share (VAS) in revenues; I5 – EBITDA margin (EBITDA to revenues).

Respondents then entered their rankings into a pairwise comparison matrix based on the principles of the AHP method.

The respondents' task was to rank the individual indicators from most to least important based on their subjective opinion. They were then asked to write the indicators in a square table or matrix, expressing the importance of a row indicator versus a column indicator using the numbers 1, 3, 5, 7 and 9 (Tabs. 1–3).

Determining the "optimal" values of profitability indicators

In the second part of the questionnaire, the experts were asked to estimate the "optimal" values for each profitability indicator as a percentage rate. Although we refer to these values as "optimal", it is acknowledged that, in the context of profitability indicators, universally optimal values cannot be strictly defined. Rather, the term "optimal" is used here to denote expert-informed target values that are considered appropriate for ensuring the long-term sustainability of businesses in this specific sector. These values reflected their opinions on what would constitute sustainable levels of profitability for businesses operating in the accommodation services sector.

After two rounds of questioning, we arrived at the final values to work with. We used a modal matrix to determine a matrix of importance values. The optimal values were determined as the median values of the indicators, as assessed by our experts. We then calculated the weights of each indicator, as shown in using three methods: the exact Saaty method, the approximate Saaty method, and the best-worst method, as shown in the following subsections. Each method has its advantages and disadvantages. From our point of view, the approximate Saaty method is probably the simplest, Damodaran (2015). However, given the availability of the Solver function in MS Office Excel, it is advisable to use the exact Saaty method, Saaty (2005). One of the most recent methods for determining weights is the best-worst method, introduced by Rezaei (2016b) and Brunelli and Rezaei (2019). The solver he constructed is available for free in Rezaei (2016b).

2.1 Saaty method

Weighting criteria is essential in decision-making when setting priorities. Several methods exist for this, notably Saaty exact method, which uses pairwise comparisons on a 9-point scale and computes weights via the eigenvector of the comparison matrix (Boďa & Úradníček, 2021; Brunelli & Rezaei, 2019).

The first step of the Saaty method is to construct a matrix that has as many rows and columns as there are indicators. We write the values from 1 to 9 in the table to express the relative importance of the row indicator compared to column one. To make the data consistent, it is obvious that we express the relative "unimportance" as the inverted value.

To verify the validity of the matrix, it is first necessary to calculate the so-called Saaty consistency index (*CI*) as follows:

$$CI = \frac{\lambda_{\max} - n}{n - 1} \tag{6}$$

where: λ_{max} – the largest positive eigenvalue of the matrix; n – number of indicators.

The so-called consistency ratio *CR* is calculated according to the formula:

$$CR = \frac{Cl}{RI} \tag{7}$$

where: *RI* – random index, which can be found in Beaver (1966) or Mu and Pereyra-Rojas (2017). For n = 5 random index *RI* = 1.11 (Altman, 1968).

For the matrix to be valid, in other words, for us to consider the matrix as sufficiently consistent, with respect to Beaver (1966) the value of *CR* must not exceed 0.10. Therefore, if the consistency ratio is at most 0.10, the matrix can be considered sufficiently consistent, otherwise it is necessary to reevaluate the mutual comparisons of criteria, or indicators.

This statement implies that the consistency index expresses the degree of inconsistency of the matrix.

Remark 1

In practice, it is rarely possible to achieve full matrix consistency. But it should not be enforced unconditionally. In the case of inconsistent matrices, an eigenvector with positive components is searched for, the sum of which is 1, corresponding to the largest eigenvalue λ_{max} of the matrix. It is appropriate that $\lambda_{max} \approx n$, and the other eigenvalues are positive and close to 0. This follows from the consideration that small changes in the matrix elements lead to only small changes in the eigenvalues.

When determining the weights, we proceed from the condition that the elements of the matrix s_{ij} differ as little as possible from the matrix v_i/v_j, and as a suitable measure we use the sum of the squares of the deviations of the elements of both matrices.

Therefore, the individual weights, which are listed in Tab. 1, can be calculated using the solver function in MS Office Excel, where the Saaty optimization criterion is minimized.

$$\min \xi = \xi^* \tag{8}$$

Under the conditions:

$$\min \sum_{i=1}^{n} \sum_{j=1}^{n} \left(s_{ij} - \frac{v_i}{v_j} \right)^2 < \xi \tag{9}$$

and $$v_1, v_2, \dots, v_n > 0 \wedge \sum_{i=1}^{n} v_i = 1 \tag{10}$$

where: s_{ij} – the individual elements of the matrix; i, j = 1, 2, …, n; v_i – the weight of the i^{th} indicator; v_j – the weight of the j^{th} indicator.

Tab. 1: Saaty exact method

i, j	**I2**	**I3**	**I1**	**I4**	**I5**	v_i
I2	1	3	5	7	9	0.4409
I3	0.3333	1	3	5	7	0.3108
I1	0.2000	0.3333	1	3	5	0.1403
I4	0.1429	0.2000	0.3333	1	3	0.0633
I5	0.1111	0.1429	0.2000	0.3333	1	0.0446
v_j	0.4409	0.3108	0.1403	0.0633	0.0446	$\sum_{i=1}^{n} v_i = 1$

Note: I1 – return on equity (ROE); I2 – return on assets (ROA); I3 – return on sales (ROS); I4 – value added share (VAS) in revenues; I5 – earnings before interest, taxes, depreciation, and amortization margin (EBITDA to revenues).

Source: own

2.2 Approximate Saaty method

The approximate Saaty method is a great help for those who do not have a solver in MS Office Excel (Boďa & Úradníček, 2021). By simply calculating the geometric mean of the values assigned to each indicator, we can determine the weights of each indicator.

Tab. 2: Approximate Saaty method

i, j	**I2**	**I3**	**I1**	**I4**	**I5**	p_i	v_i
I2	1	3	5	7	9	3.9363	0.5100
I3	0.3333	1	3	5	7	2.0362	0.2638
I1	0.2000	0.3333	1	3	5	1.0000	0.1296
I4	0.1429	0.2000	0.3333	1	3	0.4911	0.0637
I5	0.1111	0.1429	0.2000	0.3333	1	0.2541	0.0329
Sum						$\sum_{i=1}^{n} p_i = 7.7176$	$\sum_{i=1}^{n} v_i = 1$

Source: own

The individual values of p_i are determined using the following formula:

$$p_i = \sqrt[n]{\prod_{j=1}^{n} s_{ij}} \tag{11}$$

We use the formula to calculate the normalized weights from the individual p_i values as follows:

$$v_i = \frac{p_i}{\sum_{i=1}^{n} p_i} \tag{12}$$

2.3 Best-worst method

The best-worst method (BWM) was introduced by Rezaei (2015, 2016a) and involves selecting the best and worst criteria and comparing them with others, making it effective for large sets of criteria. The best-worst method is based on a sequential comparison of pairs of indicators. In addition, it is shown that over the course of several years this method has become very popular in multi-criteria decision-making due to its properties.

The method is implemented in several steps.

Step 1. A set of decision indicators is determined.

Step 2. The best and worst criterion, or indicator, is determined. If more than one criterion, or indicator, is considered the best or worst, one can be arbitrarily chosen, Rezaei (2016a). In our case, among the five indicators, the most significant one is identified – best and the least significant – worst.

Step 3. The preference for the best indicator will be gradually expressed in comparison with the other indicators using the cardinal scale, which was also used in previous methods. The value 1 again represents a match in importance. Thus, we get $A_B = (s_{B1}, ..., s_{Bn})$ as the vector of the best indicator relative to the others. In our case, among the five indicators, the most significant one is identified – best and the least significant – worst. The best in our case is indicator I2 and the worst one is I5.

Step 4. The preference for the best indicator will gradually emerge in comparison with the other indicators using the cardinal scale, which was also used in previous methods. The value 1 again represents a match in importance. Thus, we get $A_B = (s_{B1}, ..., s_{Bn})$ as the vector of the best indicator relative to the others. In our case, it is the vector $A_B = (1, 3, 5, 7, 9)$, see Tab. 3. The preference of other indicators to the worst indicator is gradually expressed by the vector $A_W = (s_{1W}, ..., s_{nW})^T$. This vector in the framework of our indicators is $A_W = (9, 7, 5, 3, 1)^T$ and is written in Tab. 3.

Step 5. Determination of the optimal weights $(v_1^*, v_2^*, ..., v_n^*)$. Optimal weights are determined using a mini-max optimization model.

$$\min \sup \left\{ \left| \frac{v_B}{v_i} - s_{Bj} \right|, \left| \frac{v_i}{v_W} - s_{iW} \right| \right\} \tag{13}$$

With conditions:

$$v_1, v_2, ..., v_n > 0 \wedge \sum_{i=1}^{n} v_i = 1 \tag{14}$$

This model is converted to the model:

$$\min \xi = \xi^* \tag{15}$$

Under conditions:

$$\left| \frac{v_B}{v_i} - s_{Bj} \right| \leq \xi \text{ and } \left| \frac{v_i}{v_W} - s_{iW} \right| \leq \xi \tag{16}$$

$$\text{and } \quad v_1, v_2, ..., v_n > 0 \wedge \sum_{i=1}^{n} v_i = 1 \tag{17}$$

By solving this model, we obtain the optimal weights $(v_1^*, v_2^*, ..., v_n^*)$.

Step 6. The last step is to calculate the level of inconsistency using a robust index called the consistency ratio CR^{BW}, which is given by:

$$CR^{BW} = \frac{\xi^*}{CI^{BW}} \tag{18}$$

where: CI^{BW} – the consistency index.

The consistency index is for the number of indicators and at the same time for the maximum value of the relative importance scale in the pairwise comparison of indicators $n = a_{BW} = 5$ at the level of 2.30 (Dijkstra, 2011;

Tab. 3: Best-worst method

	Indicator				
	I2	I3	I1	I4	I5
Selected the best	I2				
Selected the worst	I5				
The best to others A_B	1	3	5	7	9
Others to the worst A_W	9	7	5	3	1
Optimal weights v_i^*	0.5246	0.2117	0.1270	0.0907	0.0460

Source: own

Mi et al., 2019). Using the solver Rezaei (2016b), we obtained the optimal value ξ* = 0.1104, and therefore CR^{BW} = 0.0480. The consistency ratio is a number from the interval [0, 1], and the smaller it is, the more reliable the results are.

ROA (I2) is consistently the most important indicator, especially emphasized by the best-worst and approximate Saaty methods. EBITDA margin (I5) and added value in sales VAS (I4) are considered the least significant across all methods. Weighting patterns are generally consistent, though best-worst tends to compress weights more modestly than the Saaty methods. Fig. 1 shows that the best-worst method tends to assign higher weights to the most important indicator (ROA) and lower weights to the less important ones, resulting in a more polarised distribution.

In contrast, the exact and approximate Saaty methods produce more evenly distributed weights, particularly between the top three indicators (ROA, ROS and ROE). This indicates a more balanced view of importance. The exact and approximate Saaty methods yield very similar results, reflecting the fact that the approximation does not significantly deviate from the exact calculation. The best-worst method emphasizes extremes more strongly, whereas the Saaty methods distribute importance more evenly.

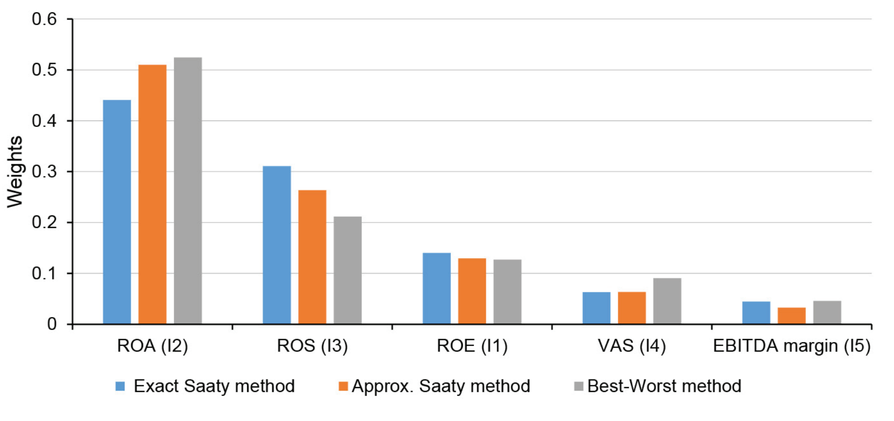

Fig. 1: Weights of individual profitability indicators

Source: own

The resulting indicator weights provide the basis for constructing a composite profitability index.

3 Results and discussion: Overall profitability index

This chapter presents the main findings of our investigation into selected profitability indicators. It recalls the procedure and data processing, and presents the results; specifically, the overall profitability index of accommodation facilities in Slovakia as determined in this study.

To reach a consensus on the "optimal" values, it was necessary to conduct a multi-round survey using a modified Delphi method. In this iterative process, experts were provided with anonymous summaries of the results from previous rounds, allowing them to revise their responses if they wished. The responses in the final round showed no significant deviations, and the final "optimal" values used in the construction of the profitability index were derived as the median of the experts' estimates. These values are reported in Tab. 4. The resulting weight coefficients and "optimal" profitability values formed the basis for constructing the profitability index for enterprises in the accommodation services sector. The index reflects not only the financial indicators themselves but also the expert consensus on their relative importance and target ("optimal") values, thereby providing a more context-sensitive tool tailored to this specific segment of the tourism industry.

Based on the optimal values of the individual indicators and the sectoral median values CRIBIS.SK (2025), we calculated the initial values of the indicators under study on a 100-point scale. All scores were determined within a linear relationship, as illustrated for the ROE indicator in Fig. 2.

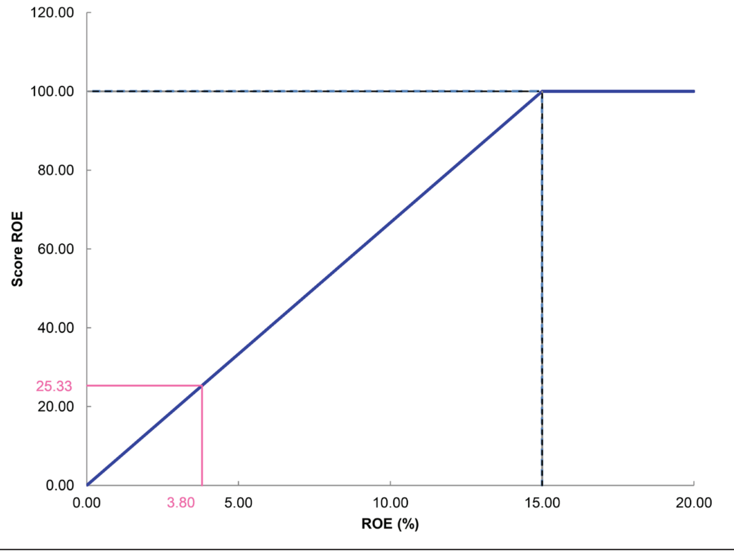

Fig. 2: Score of the profitability indicator ROE (I2) for year 2023

Note: The values displayed in purple – real value ROE (I1) and corresponding score for year 2023.

Source: own

The optimal value of each indicator is determined based on the responses of our experts and is always assigned a score of 100. A score of 0 is assigned to each indicator with a zero value.

The corresponding scores are reported in Tab. 4.

To improve orientation regarding the growing development of the data in Tab. 4,

Tab. 4: Scores of individual profitability indicators

	I2	I3	I1	I4	I5
Optimal value (median; %)	5	6	15	30	15
Relevant ratio					
CRIBIS.SK 2020 score*	−0.16 0.00	−0.26 0.00	−1.47 0.00	20.30 67.67	3.52 23.47
CRIBIS.SK 2021 score*	−0.01 0.00	−0.06 0.00	−0.14 0.00	17.36 57.87	4.45 29.67
CRIBIS.SK 2022 score*	0.01 0.20	0.00 0.00	2.35 15.67	23.36 77.87	6.83 45.53
CRIBIS.SK 2023 score*	1.02 20.40	0.79 13.17	3.80 25.33	25.59 85.30	8.43 56.20

Note: I1 – return on equity (ROE); I2 – return on assets (ROA); I3 – return on sales (ROS); I4 – value added share (VAS) in revenues; I5 – earnings before interest, taxes, depreciation, and amortization margin (EBITDA to revenues).
* row 1 – real values of the indicators from CRIBIS.SK; row 2 – the indicators scores directly proportional to the optimal values, which have a score of 100 points.

Source: own

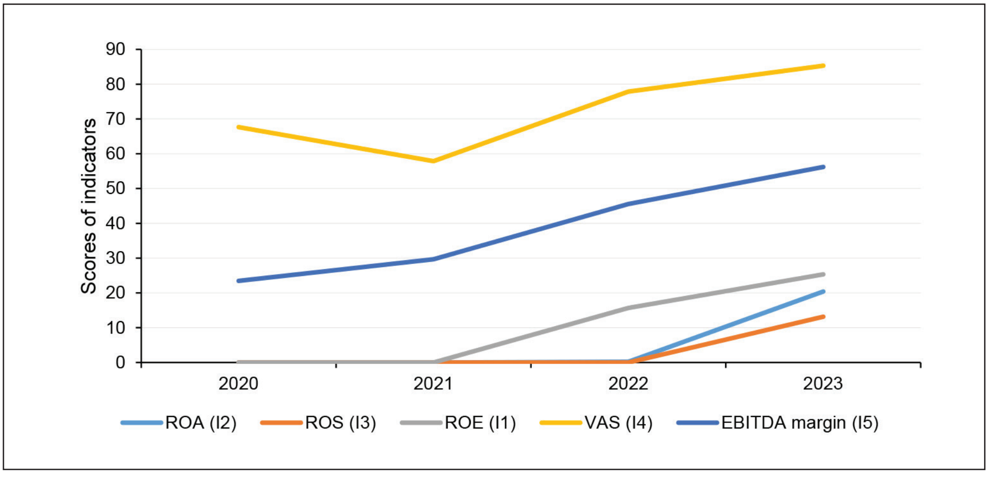

Fig. 3: The growth trend of the examined indicators scores on a scale of 100 points

Source: own

we also present a graphical representation of the development of the indicators examined on Fig. 3.

The overall profitability index for each year is calculated as the weighted arithmetic mean of the indicator scores, using normalised weights determined by the exact Saaty method, by the approximate Saaty method or the best-worst method as follows, see Tab. 5.

$$WAM = \sum_{i=1}^{5} v_i \times \text{score } I_i \qquad (19)$$

All methods show consistently increasing trends, reflecting strong post-COVID recovery. The best-worst method yields the highest profitability scores each year. The exact and

 E+M 2026, volume 29, issue 2, pp. 92–106, DOI: 10.15240/tul/001/2026-2-007

Tab. 5: The overall profitability index (2020–2023)

Year	Overall profitability index			
	Exact Saaty method	Approximate Saaty method	Best-worst method	Average score (individual years)
2020	5.34	5.08	7.22	5.88
2021	4.99	4.66	6.61	5.42
2022	9.25	8.59	11.25	9.70
2023	24.56	24.44	27.03	25.34
Score range (2020–2023)	19.56	19.78	20.42	

Source: own

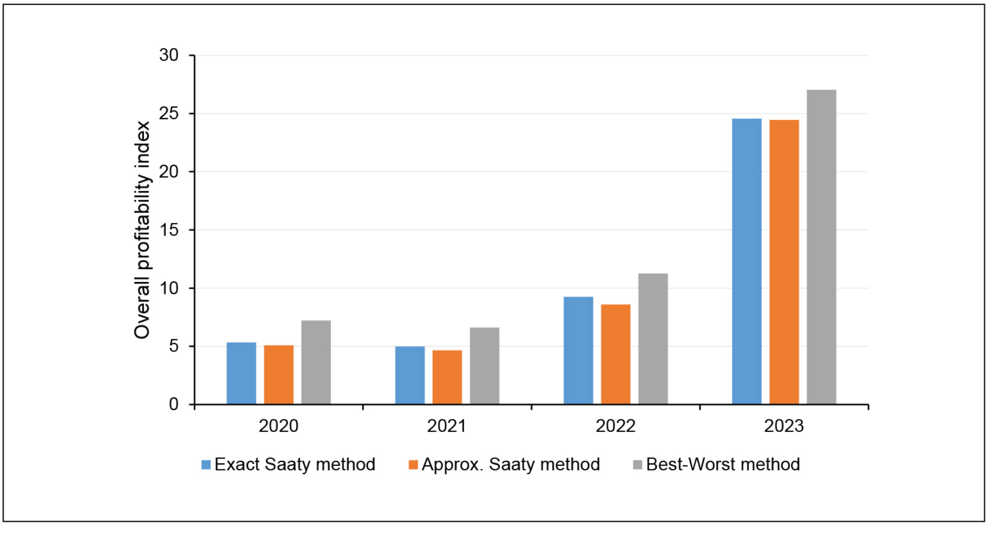

Fig. 4: The overall profitability index

Source: own

approximate Saaty methods closely track each other, especially in 2023 (Fig. 4).

The results provide a response to the formulated research question. The indicators ROA, ROS, and ROE emerged as the most significant, with ROA receiving the highest weight across all applied methods (exact Saaty, approximate Saaty, and best-worst). The value of the composite index varied depending on the weighting method: the best-worst method yielded the highest score, whereas the Saaty methods produced a more balanced distribution of weights. These findings confirm that the composite index enables a more comprehensive assessment of financial performance than individual indicators.

Comparison with the literature indicates that our results are consistent with Santos et al. (2022), who emphasize the importance of ROS; however, in our study, ROA emerged as the most critical indicator. Similarly, as noted by Maxim (2023) regarding the relevance of the EBITDA margin, our findings also underscore its significance, albeit with a lower weight. Tangen (2005) and Costa and Costa (2024) highlight the need for a broader perspective on value creation, which is reflected in our composite index through the inclusion of the VAS indicator.

Moreover, the differences observed between the weighting methods suggest distinct approaches to indicator prioritization. The best-worst method accentuates extremes, which may be suitable in contexts, where identifying a dominant factor (e.g., ROA) is critical, whereas the Saaty methods provide a more balanced perspective, appropriate for complex decision-making. This distinction has practical implications: firms with high capital intensity may favor an approach emphasizing asset utilization efficiency. Furthermore, the results indicate that incorporating broader indicators, such as VAS, can better capture a firm's value-creation capability, which is particularly important during periods of economic turbulence.

Conclusions

The accommodation industry is currently navigating a period of recovery following a series of significant external shocks, including the COVID-19 pandemic, the energy crisis, and persistently high inflation. These disruptions have revealed structural vulnerabilities within the sector and highlighted the need for greater resilience, adaptability, and financial prudence. In response, accommodation providers are increasingly focused on eliminating operational inefficiencies, fostering continuous improvement, and leveraging technological innovations that streamline processes and enhance the overall guest experience. In this context of heightened uncertainty and economic pressure, ensuring long-term financial sustainability is paramount. Ensuring long-term financial sustainability in this context requires a more comprehensive approach to performance evaluation, moving beyond isolated profitability indicators toward composite indices that capture the multidimensional nature of financial health. Such indices facilitate more accurate benchmarking, allow analysis of performance trends over time, and provide strategic insights supporting informed decision-making. By integrating operational optimization with sophisticated financial performance monitoring, the sector can strengthen its competitive positioning and build resilience against future disruptions.

Although the proposed composite index has not yet been applied at the firm level, its calculation using sectoral data (CRIBIS.sk) demonstrates its practical applicability. The study has some inherent limitations typical for methodological research of this type. The expert panel comprised 14 respondents, which, while sufficient to ensure methodological rigor, may limit generalizability. Enterprise-level case studies were not included, which provides an opportunity for future validation at the firm level. Additionally, the use of median-based "optimal" indicator values offers a pragmatic approach to standardization, though future studies could explore more granular methods to capture sectoral heterogeneity. The analysis covers the period 2020–2023, providing insight into post-pandemic recovery, while longer-term trends remain to be explored. Future research could build on these foundations by expanding the expert panel, incorporating longitudinal data, and applying the index to individual accommodation facilities to further validate its effectiveness in identifying financial strengths and weaknesses and supporting strategic decision-making. Advanced weighting techniques, such as the fuzzy best-worst method with triangular or intuitionistic fuzzy numbers (Cheng & Chen, 2024), could also enhance the robustness and adaptability of the index in uncertain environments.

Acknowledgments: *Jana Spirkova was supported by grant No. 1/0124/24 "Slovakia in the Context of the Pension Index and Population Ageing Index – Future Perspective" of the Slovak Scientific Grant Agency VEGA. Miroslava Vinczeova was supported by grant No. 1/0229/25 "Eco-Innovation Activities of Small and Medium-Sized Enterprises and Their Importance in Achieving the Sustainable Development Goals in the Context of the 2030 Agenda" of the Slovak Scientific Grant Agency VEGA.*

References

Alrawabdeh, W. (2021). Seasonal balancing of revenue and demand in hotel industry: The case of Dubai City. *Journal of Revenue and Pricing Management*, *21*(1), 36–49. https://doi.org/10.1057/s41272-021-00290-6

Altman, E. I. (1968). Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *The Journal of Finance*, *23*(4), 589–609. https://doi.org/10.2307/2978933

Aqil, M., Ahmed, R., Vveinhardt, J., & Streimikiene, D. (2019). Factors influencing the profitability of heavy vehicle industry: A case of Pakistan factors influencing the profitability of heavy vehicle industry: A case of Pakistan. *Montenegrin Journal of Economics*, *15*(1), 61–72. https://doi.org/10.14254/1800-5845/2019.15-1.5

Beaver, W. H. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research*, *4*, 71–111. https://doi.org/10.2307/2490171

Blahušiaková, M. (2021). Impact of COVID-19 crises on accounting entities providing accommodation services in Slovakia. In *Proceedings of 13th International Scientific Conference: Financial Management of Firms and Financial Institutions* (pp. 7–16). Faculty of Economics, VSB – Technical University of Ostrava.

Boďa, M., & Úradníček, V. (2021). *Stochastická analýza neistoty vo váhach významnosti v rámci viacaspektného hodnotenia* [Stochastic analysis of uncertainty in significance weights in a multi-aspect evaluation]. Belianum. Matej Bel University in Banská Bystrica.

Bouwens, J., de Kok, T., & Verriest, A. (2019). The prevalence and validity of EBITDA as a performance measure. *Accounting Auditing Control*, *25*(1), 55–105. https://doi.org/10.3917/cca.251.0055

Brunelli, M., & Rezaei, J. (2019). A multiplicative best-worst method for multi-criteria decision making. *Operations Research Letters*, *47*(1), 12–15. https://doi.org/10.1016/j.orl.2018.11.008

Cheng, X., & Chen, C. (2024). Decision making with intuitionistic fuzzy best-worst method. *Expert Systems with Applications*, *237*(A), 121215. https://doi.org/10.1016/j.eswa.2023.121215

Costa, V., & Costa, C. (2024). Tourism, economic development and the corporate performance of the hospitality industry: An empirical study of Portuguese tourism regions. *European Journal of Tourism Research*, *38*, 3816. https://doi.org/10.54055/ejtr.v38i.3622

CRIBIS.SK. (2025). *Univerzálny register. Odvetvové pomerové ukazovatele: 55 – Ubytovanie* [Universal register. Sectoral ratio indicators: 55 – Accommodation]. https://www.cribis.sk/

Damodaran, A. (2015). *Applied corporate finance*. John Wiley & Sons, Inc.

Dijkstra, T. K. (2011). On the extraction of weights from pairwise comparison matrices. *Central European Journal of Operations Research*, *21*(1), 103–123. https://doi.org/10.1007/s10100-011-0212-9

D'Souza, J., Ramesh, K., & Shen, M. (2010). Disclosure of GAAP line items in earnings announcements. *Review of Accounting Studies*, *15*(1), 179–210. https://doi.org/10.1007/s11142-009-9100-0

Elexa, Ľ. (2021). *Vplyv legislatívnych faktorov na finančnú výkonnosť malých a stredných podnikov* [The impact of legislative factors on the financial performance of SMEs]. Belianum. Matej Bel University in Banská Bystrica.

Enz, C. A. (2009). *Hospitality strategic management: Concepts and cases* (2nd ed.). John Wiley & Sons, Inc.

Fadhillah, I. R., Nadiyah, N., Rohmah, L., Haryadi, D., & Wahyudi, W. (2024). Profitability ratio analysis to assess the financial performance. *International Journal of Applied Finance and Business Studies*, *11*(4), 964–969. https://doi.org/10.35335/ijafibs.v11i4.243

Flesher, D. L., & Previts, G. J. (2013). Donaldson Brown (1885–1965): The power of an individual and his ideas over time. *Accounting Historians Journal*, *40*(1), 79–101. https://doi.org/10.2308/0148-4184.40.1.79

Ishizaka, A., & Labib, A. (2011). Review of the main developments in the analytic hierarchy process. *Expert Systems with Applications*, *38*(11), 14336–14345. https://doi.org/10.1016/j.eswa.2011.04.143

Jović, Z., & Tomašević, S. (2021). Traditional vs modern approaches to measuring the performance of a company. In *Proceedings of the 8th International Scientific Conference – FINIZ 2021* (pp. 31–37). Singidunum University. https://doi.org/10.15308/finiz-2021-31-37

Jugović, T., & Maričić, M. (2024). How did hotels measure performance during Covid-19? Insights from Central and Eastern Europe. *Business Systems Research Journal*, *15*(1), 201–225. https://doi.org/10.2478/bsrj-2024-0010

Lee, S. (2024). Technological innovation and profit. *Journal of Open Innovation: Technology, Market, and Complexity*, *10*(3), 100335. https://doi.org/10.1016/j.joitmc.2024.100335

Lesáková, Ľ., Ondrušová, A., & Vinczeová, M. (2019). Factors determining profitability of small and medium enterprises in selected industry of mechanical engineering in the Slovak

Republic – The empirical study. *E&M Economics and Management*, *22*(2), 144–160. https://doi.org/10.15240/tul/001/2019-2-010

Maxim, L. G. (2023). EBITDA, one of the most important indicators regarding the financial sustainability of retail companies. In *Theoretical and applied economics. Economic convergence in European Union (ECEU) Transforming the future of work: Navigating the intersection of digital transformation and human capital* (pp. 134–143). Bucharest University of Economic Studies.

Mi, X., Tang, M., Liao, H., Shen, W., & Lev, B. (2019). The state-of-the-art survey on integrations and applications of the best worst method in decision making: Why, what, what for and what's next? *Omega*, *87*, 205–225. https://doi.org/10.1016/j.omega.2019.01.009

Mu, E., & Pereyra-Rojas, M. (2017). *Practical decision making: An introduction to the analytic hierarchy process (AHP) using super decisions V2*. Springer International Publishing.

Mundi, H. S., Gautam, J., & Vasisht, S. (2023). Determinants of capital structure in the hospitality industry. *The Journal of Hospitality Financial Management*, *31*(1), 19–31. https://doi.org/10.7275/z6gp-an42

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research*, *18*(1), 109–131. https://doi.org/10.2307/2490395

Palepu, K. G., Healy, P. M., & Peek, E. (2022). *Business analysis and valuation* (IFRS). Cengage Learning EMEA.

Rezaei, J. (2015). Best-worst multi-criteria decision-making method. *Omega*, *53*, 49–57. https://doi.org/10.1016/j.omega.2014.11.009

Rezaei, J. (2016a). Best-worst multi-criteria decision-making method: Some properties and a linear model. *Omega*, *64*, 126–130. https://doi.org/10.1016/j.omega.2015.12.001

Rezaei, J. (2016b). *BWM-Solver-4.xlsx* [software]. https://bestworstmethod.com/software/

Saaty, T. L. (1977). A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, *15*(3), 234–281. https://doi.org/10.1016/0022-2496(77)90033-5

Saaty, T. L. (2005). *Theory and applications of the analytic network process: Decision making with benefits, opportunities, costs, and risks*. RWS Publication.

Santos, L. L., Gomes, C., Malheiros, C., & Campos, F. (2022). Measuring the hotels' performance using profitability ratios. *Academy of Accounting and Financial Studies Journal*, *26*(6), 1–12.

Šebová, Ľ., Sokologorsky, T., & Marčeková, R. (2024). The current state of business environment in Slovakia and its impact on tourism enterprises. In V. Katsoni & G. Cassar (Eds), *Recent advancements in tourism business, technology and social sciences* (pp. 297–312). Springer Nature Switzerland. https://doi.org/0.1007/978-3-031-54342-5_18

Setyopurnomo, R., Wiryono, S. K., Lestari, Y. D., & Sukarno, S. (2025). Enhancing operation effectiveness by maximizing daily EBITDA as a turnaround strategy execution system to enhance operation performance and profitability. *Journal of Open Innovation: Technology, Market, and Complexity*, *11*(1), 100431. https://doi.org/10.1016/j.joitmc.2024.100431

Tangen, S. (2005). Demystifying productivity and performance. *International Journal of Productivity and Performance Management*, *54*(1), 34–46. https://doi.org/10.1108/17410400510571437

Vojteková, S., & Klieštik, T. (2024). The impact of the COVID-19 pandemic on profitability indicators in the hospitality sector. *Management Dynamics in the Knowledge Economy*, *12*(1), 39–53. https://doi.org/10.2478/mdke-2024-0003

Green innovation and firm performance in China: Empirical insights from A-share listed manufacturers

Ganglong Wu[1], Syed Ahmed Salman[2], Baoting Huang[3], Tengkun Wang[4]

[1] Liming Vocational University, School of Business, China, ORCID: 0009-0000-5503-0176, ganglong@lincoln.edu.my (corresponding author);
[2] Lincoln University College, School of Business and Management, Malaysia, ORCID: 0000-0001-7101-5961, syedahmed@lincoln.edu.my;
[3] Liming Vocational University, School of Business, China, ORCID: 0000-0001-7456-0842, huang_baoting2732@163.com;
[4] Liming Vocational University, School of Business, China, ORCID: 0009-0008-7102-8613, timwang@lmu.edu.cn.

Abstract: *As environmental challenges intensify globally, green innovation has become a strategic imperative for firms seeking to balance ecological responsibility with sustained competitiveness. We examine whether green innovation improves firm performance in China's manufacturing sector. Using panel data on 1,033 A-share manufacturers (2014–2023; 10,330 firm-years), we estimate two-way fixed-effects models (firm and year fixed effects (FE)) with cluster-robust standard errors (SE). Green innovation is measured by green patent outputs; firm performance is proxied by average return on equity (ROEAVG), with return on invested capital (ROIC) as a robustness outcome. Results show a positive and statistically significant association on average, with stronger effects in more developed regions. Findings remain under lagged-specification checks and alternative outcome measures. Conceptually, we extend the resource-based view by framing green innovation as a capability bundle whose payoff is conditioned by institutional context and absorptive capacity. Practically, predictable enforcement and disclosure raise both the quantity and influence of green patents and improve market pricing of eco-innovation; targeted, adoption-cost-reducing incentives tied to substantive innovation help curb greenwashing.*

Keywords: *Green innovation, firm performance, resource-based view, manufacturing sector, regional heterogeneity, sustainable development.*

JEL Classification: *D22, L20, O30, M14, Q01.*

APA Style Citation: Wu, G., Salman, S. A., Huang, B., & Wang, T. (2026). Green innovation and firm performance in China: Empirical insights from A-share listed manufacturers. *E&M Economics and Management*, *29*(2), 107–123. https://doi.org/10.15240/tul/001/2026-2-008

Introduction

In recent years, the intensifying global environmental crisis, marked by climate change, ecological degradation, and resource scarcity, has propelled green development to the forefront of policy and business agendas (Ostapenko et al., 2024). As a strategic response, green innovation, defined as the implementation of environmentally friendly processes, products, and management practices, has emerged as a core pillar of sustainable enterprise transformation (Levchenko et al., 2021). It not only reflects a firm's commitment to corporate social responsibility but also serves as a source of long-term competitive advantage in an increasingly eco-conscious market (Tian et al., 2023; Zheng & Iatridis, 2022).

In China, the central government has integrated green development into national policy priorities, creating an institutional environment conducive to green innovation through subsidies, regulatory mandates, and green financing mechanisms (Bai et al., 2019). These policy efforts have accelerated the integration of green innovation within manufacturing sectors, particularly in regions with advanced industrial infrastructure. Consequently, understanding the impact of green innovation on firm performance is not only of academic interest but also of practical relevance for managers and policymakers seeking to align economic and environmental goals.

Although green innovation is widely acknowledged as a critical driver of sustainable development, empirical evidence regarding its impact on firm performance remains inconsistent and unsettled. Some researchers contend that it boosts profitability through enhanced resource utilization, waste minimization, and the creation of novel market avenues (Valaskova et al., 2022). From the perspective of RBV, green innovation constitutes a rare, valuable, and inimitable capability that can generate sustainable performance differentials (Barney et al., 2021). Conversely, other studies highlight the uncertain returns and high investment risks associated with green R&D, especially for firms in less developed regions or with limited absorptive capacity (Levchenko et al., 2021).

Given this ambiguity, there is a need for a more nuanced investigation that accounts for contextual factors such as regional heterogeneity, firm characteristics, and policy environments. This research responds to that need by analyzing panel data from 1,033 manufacturing firms listed on the A-share listed market in China, including the Shanghai and Shenzhen exchanges, during the period from 2014 to 2023, aiming to empirically explore the linkage between green innovation and business performance. Specifically, the study aims to: (1) test whether green innovation improves firm performance; and (2) explore whether this effect varies across regional contexts.

This study contributes to the literature by offering context-sensitive insights into the performance implications of green innovation and extending the resource-based view in the sustainability context.

1 Literature review

1.1 Definition and dimensions of green innovation

Green innovation, often termed environmental or eco-innovation, encompasses "new or significantly improved products (goods or services), processes, marketing methods, organizational structures or institutional arrangements that result in environmental benefits" (Shouwen Wang et al., 2022; Vimalnath et al., 2022). It extends beyond mere technological upgrades and involves rethinking business models and managerial practices that reduce environmental burdens (Abadzhiev et al., 2022).

Scholars typically classify green innovation into three dimensions: (1) green product innovation, which involves the development of environmentally friendly goods and services; (2) green process innovation, which refers to the adoption of cleaner production methods and resource-efficient technologies; and (3) green managerial or organizational innovation, which entails institutionalizing sustainability practices within organizational routines.

1.2 Green innovation and firm performance: Empirical evidence

A growing body of empirical research investigates the relationship between green innovation and firm performance. Several studies affirm the positive financial implications of green innovation, asserting that green innovation can improve cost efficiency, attract eco-conscious customers, and differentiate products in saturated markets (Tian et al., 2023; Zheng & Iatridis, 2022). For example, Valaskova et al. (2022) found that green practices significantly enhance return on assets across European firms, particularly in capital-intensive sectors.

Across 160 studies and 124,000 firms, Zheng and Iatridis (2022) report positive average impacts, with managerial eco-innovation often yielding stronger payoffs than purely technological upgrades. Firm-level evidence in Chinese A-shares echoes this nuance: high-quality green innovation boosts corporate financial performance, whereas low-quality activity may not (L. Liu et al., 2024). In heavily polluting sectors, accumulated green patent stocks are priced by the market (Tian et al., 2023).

However, not all studies confirm a uniformly positive relationship. Some suggest a non-linear or uncertain impact, emphasizing the initial cost burden, longer return cycles, and potential trade-offs between green and traditional innovation priorities (Mirza et al., 2025). These mixed findings underscore the importance of institutional context, regional capabilities, and firm absorptive capacity in shaping outcomes.

1.3 Gaps in existing research and theoretical motivation

Yet these syntheses also expose unresolved gaps. Results hinge on what we measure and where we look. This calls for context-aware tests and measurement-robust designs in the Chinese setting.

While the literature affirms the theoretical potential of green innovation, several gaps remain. First, most studies focus on developed economies or high-tech firms, offering limited exploration of emerging markets like China, where regulatory environments and market dynamics differ significantly. Second, many analyses conceptualize green innovation as a binary construct, overlooking its heterogeneity across regions, industries, and innovation types. Third, there is limited integration between strategic management theory, such as the resource-based view, and the empirical testing of green innovation's impact.

Given these gaps, this study aims to enrich the discourse by combining resource-based theory with regionally disaggregated data from Chinese manufacturing firms to explore how green innovation affects firm performance and whether this relationship is moderated by regional contexts.

2 Theoretical analysis and research hypotheses

2.1 Resource-based view and the strategic role of green innovation

Resource-based view (RBV), one of the most influential paradigms in strategic management, posits that firms achieve sustainable competitive advantage by cultivating resources and capabilities that are valuable, rare, inimitable, and non-substitutable (Barney et al., 2021). Intangible assets, including technological know-how, organizational routines, and innovation capacity, are particularly emphasized in this theory.

Green innovation, defined as the development and implementation of products, processes, or management practices that significantly reduce environmental harm, is increasingly viewed as a form of strategic intangible asset (Cheng et al., 2025). It enhances resource efficiency, reduces regulatory and reputational risks, and can serve as a barrier to entry for competitors. Moreover, it can support firms in gaining access to green financing and preferred supply chains, particularly in sustainability-conscious markets (Singh et al., 2020).

In this theoretical context, green innovation can be conceptualized as a dynamic capability that enables firms to respond to both environmental turbulence and stakeholder pressure (Huang & Xiao, 2023). It is not merely a symbolic gesture but a core transformation in resource deployment, organizational alignment, and performance orientation.

2.2 Knowledge gaps and conceptual contention

Although many studies acknowledge the potential of green innovation to generate long-term returns, the empirical evidence remains mixed. Some scholars argue that green innovation can improve financial performance by reducing input costs, unlocking new markets, and differentiating products (Tian et al., 2023; Zheng & Iatridis, 2022). Others contend that green innovation often incurs high up-front investment, uncertain returns, and increased complexity in R&D and production, which may not translate into immediate performance gains (Maihaiti et al., 2025).

Furthermore, existing research often overlooks how institutional factors, such as regional regulatory intensity, industrial development, and local market sophistication, moderate the impact of green innovation.

2.3 Theoretical consolidation and mechanism unpacking

Anchored in the resource-based view (RBV) and dynamic capabilities theory, we conceptualize green innovation as a capability bundle that reconfigures routines to deliver three classes of payoffs process eco-efficiency (cost reduction), green product differentiation (revenue enhancement), and compliance and reputation (risk mitigation). This perspective explains why identical green inputs can yield different performance elasticities: firms with stronger absorptive capacity and integration routines monetize eco-innovation more effectively (Zheng & Iatridis, 2022). Context also matters: in heavily polluting sectors, accumulated stocks of green technologies are recognized and priced by capital markets, reinforcing the business case for substantive green innovation (Tian et al., 2023). Complementing RBV, signaling andlegitimacy arguments clarify why predictable enforcement and disclosure increase both the quantity and the influence of green patents, thereby improving market pricing of eco-innovation. These institutional features enhance the credibility and diffusion of substantive efforts and are consistent with the stronger effects we observe in regions with tighter enforcement and greater market maturity.

2.4 Research questions and hypotheses

Building on the above discussion, this study seeks to address two core research questions:

RQ1: Does green innovation contribute positively to firm-level financial performance in the Chinese manufacturing sector?

RQ2: To what extent does green innovation influence firm performance differently across regions characterized by varying institutional frameworks and economic conditions?

To empirically explore these two questions, we anchor our reasoning in the resource-based view (RBV) which frames green innovation as a valuable, firm-specific strategic resource and derive the following testable hypotheses:

H1: Enterprise green innovation has a positive and significant effect on firm performance.

This hypothesis is grounded in the notion that green innovation enables firms to reduce resource consumption, improve production efficiency, and enhance market competitiveness and brand value through the development of environmentally friendly products and processes. As a strategic resource of firms, green innovation can significantly improve their financial performance and enhance their market position. Thus, we hypothesize that green innovation positively and significantly affects firm performance.

H2: There is regional heterogeneity in the impact of green innovation on firm performance.

Although green innovation is generally theorized to improve firm performance, its impact may vary significantly across regions. Such regional heterogeneity may stem from differences in local policies, market demand, and levels of technological advancement. For example, some regions may focus more on the implementation of environmental regulations, and enterprises in these regions may receive more policy support for green innovation and thus perform more prominently in terms of performance. Conversely, in regions where green technology adoption and environmental awareness lag behind, the performance-enhancing effects of green innovation may be less pronounced.

3 Research methodology

3.1 Variable selection and description

Following established practice in the innovation-performance literature, this study specifies firm performance as the dependent variable, green innovation as the focal independent variable, and a set of firm-level controls to account for confounding influences.

Regarding the measurement of firm performance, existing studies have primarily used return on assets, return on equity, net profit margin on total assets, Tobin’s Q value (de Oliveira & Basso, 2024; M. Liu et al., 2024). Some scholars combine return on total and net assets as composite indicators of firm performance (Sadiq & Gebba, 2022). Based on the aforementioned literature, average return on equity (ROEAVG) is selected as the primary measure of firm performance. Meanwhile, return on invested capital (ROIC) is selected as a robustness test proxy variable. All outcomes are estimated under the same design.

Green innovation refers to enterprise activities aimed at reducing resource consumption, lowering emissions, and enhancing environmental outcomes through technological, product, or process innovation (Cheng et al., 2025). In existing studies, the main variables are the number of green patent applications, the number of citations to green patents, and the number of green patents granted (Block et al., 2025; Do, 2024; Liu et al., 2025; B. Wang et al., 2021) are used as indicators to measure the green innovation capability of an enterprise. Drawing on prior studies and based on data accessibility, green innovation is measured by summing the number of green invention patents and green utility model patents independently filed by listed firms in a given year. While green patent counts are widely used in the literature to quantify firms’ green innovation activities (Block et al., 2025; Do, 2024; Liu et al., 2025; B. Wang et al., 2021), this measure primarily captures technological outputs and may underrepresent non-patented, tacit, or managerial innovations such as process optimization and organizational restructuring. As highlighted by Kemp et al. (2023), many eco-innovations are not patentable but still contribute substantially to environmental performance. Similarly, the OECD's Eco-Innovation Indicators recommend complementing patent-based measures with survey or performance-based data. Given the availability and comparability of patent statistics in China, this study adopts number of independent applications for green patents as the primary indicator but explicitly recognizes its limitations. To mitigate the effect of zero values in the dataset, one is added to the total patent count for each year during the observation period, and the natural logarithm of this adjusted value is used as the final indicator.

The accurate identification of control variables plays a vital role in establishing a sound analytical framework for linking sustainability-oriented innovation with firm-level performance indicators. Firm size, R&D investment, firm age, and asset-liability ratio are key determinants of innovation capability and performance, financial stability, and responsiveness to market changes. The establishment time of an enterprise affects its experience, resource accumulation, and ability to respond to market changes. The asset-liability ratio affects the financial stability and investment capacity of an enterprise. Considering data availability, enterprise size, enterprise age, asset-liability ratio, and R&D expense investment are selected as control variables (Meng et al., 2020; Wang et al., 2024; Zhang et al., 2022). In order to avoid linearized non-linear relationships, heteroskedasticity, and to reduce the effect of data size, as well as to better capture the relationship between the variables, logarithmic treatment was taken for the above control variables.

The meaning of each variable and the description of its value are shown in Tab. 1.

Tab. 1: Meaning of variables and description of values

Type	Variable name	Variable code	Variable value description
Dependent variable	Firm performance	*ROEAVG*	Net income divided by average shareholders' equity
Independent variable	Green innovation	*GI*	LN(number of independent applications for green patents in the current year + 1)
Control variables	Enterprise size	*ASSETS*	LN(balance of total assets at the end of the current year)
	Enterprise age	*AGE*	LN(current year − year of enterprise establishment)
	Asset-liability ratio	*ALR*	LN(ratio of total liabilities to total assets)
	R&D expenditure	*R&D*	LN(R&D expenditure as a percentage of operating revenue + 1)

Source: own

3.2 Data source and sample selection

This research focuses on A-share manufacturing firms listed on the Shanghai and Shenzhen stock exchanges during the period from 2014 to 2023. Firms labeled as ST or *ST and those with missing data during the research period were excluded. The final dataset comprises 1,033 firms, yielding 10,330 firm-year observations. Data on green patent applications and citations were collected from a specialized

Chinese research data platform, whereas other variables were sourced from the Oriental Fortune Choice database. To overcome the problem that the exported ROEAVG data may have centralized extreme values or outliers, which could affect the statistical analysis results, we used winsorize to perform 1% and 99% percentile tailing on the data the ROEAVG variable.

3.3 Model construction

Following Baltagi (2021), we select the two-way fixed-effects model (firm and year FE) suitable for large-*N*, short-*T* panel data for two core reasons: (1) firm FE controls for time-invariant unobserved heterogeneity at the firm level (e.g., inherent innovation capabilities, organizational culture, and long-term governance structures), which is critical given the 1,033 diverse manufacturing firms in the sample; (2) year FE absorbs macro-level time trends (e.g., national environmental policies like China's "12th Five-Year Plan" (2011–2015), "13th Five-Year Plan" (2016–2020) and "14th Five-Year Plan" (2021–2025), as well as economic cycles). This aligns with standard practices in firm-level innovation-performance studies (L. Liu et al., 2024) to mitigate omitted variable bias.

For the avoidance of ambiguity, we formally define the "two-way fixed-effects (two-way FE) model" adopted in this study as: a panel data estimation framework that simultaneously controls for firm fixed effects (to absorb time-invariant firm-specific heterogeneity, such as inherent innovation capabilities or long-term environmental strategy) and year fixed effects (to absorb macro-level time trends, such as national environmental policies or economic cycles). All subsequent regression analyses (including baseline regressions, regional heterogeneity tests, and robustness checks) strictly follow this specification.

To test *H1* and *H2*, a two-way fixed-effects model was constructed.

$$\begin{aligned} ROEAVG_{i,t} &= \beta_0 + \beta_1 GI_{i,t} + \beta_2 ASSETS_{i,t} + \\ &+ \beta_3 AGE_{i,t} + \beta_4 ALR_{i,t} + \beta_5 R\&D_{i,t} + \\ &+ u_i + t_t + \varepsilon_{i,t} \end{aligned} \tag{1}$$

$$\begin{aligned} ROEAVG(region)_{i,t} &= \beta_0 + \beta_1 GI(region)_{i,t} + \\ &+ \beta_2 ASSETS_{i,t} + \beta_3 AGE_{i,t} + \beta_4 ALR_{i,t} + \\ &+ \beta_5 R\&D_{i,t} + u_i + t_t + \varepsilon_{i,t} \end{aligned} \tag{2}$$

where: β_0 – the intercept; β_1 – the regression coefficient of green innovation of the enterprise; β_2 – the coefficient of the control variable enterprise size; β_3 – the coefficient of the control variable enterprise age; β_4 – the coefficient of the control variable asset liability ratio; β_5 – the control variable coefficient of enterprise R&D expenditure investment; μ_i – firms fixed effect; t_t – year fixed effect; ε – the error term; *i* – manufacturing firms listed on the Shanghai and Shenzhen A-share markets in China, and *t* – the year.

Hypotheses 1–2 are supported if the estimated coefficient β_1 is positive and statistically significant, otherwise, it is not supported.

3.4 Assumptions and identification notes

We maintain two baseline hypotheses and avoid mediator and moderator proliferation to preserve parsimony. Identification rests on three design choices already embedded in our empirical strategy: (1) temporal ordering via lagged green innovation measures to reduce simultaneity; (2) comprehensive controls (size, age, asset-liability, R&D) and firm and year fixed effects to absorb structural differences; (3) indicator transparency reported in robustness checks. Given evidence that indicator choice affects inference (Favot et al., 2023) and that institutional enforcement shifts both the quantity and influence of green patents (He & Qiu, 2025), we interpret coefficients as average effects under stated assumptions, and discuss non-tested contingencies in the conclusion.

4 Results

4.1 Summary description of key metrics

Tab. 2 displays the statistical summaries of the key variables employed in this research. Firm performance (ROEAVG) ranges from −44.2021 to 34.5801, with a standard deviation of 11.0773, indicating that there are certain differences in the profitability of the sample companies during the study period, but overall, it is relatively reasonable. The mean and median values are 6.1737 and 6.3090, respectively, suggesting moderate overall profitability across the sample. The mean value of green patent applications of enterprises after taking logarithmic treatment is 0.4679. The median and minimum values are both zero, suggesting that the overall level of green innovation among sampled firms is relatively low. All other control variables are logarithmically processed in the early stage, and their mean, standard deviation, maximum value, minimum value and median fall within acceptable ranges, showing no presence of outliers or anomalies.

Tab. 2: Summary description of key metrics (original and natural logarithm)

Variable		Obs.	Mean	Std. dev.	Min	Max	Median
Natural logarithm variable	*ROEAVG*	10,330	6.1737	11.0773	–44.2021	34.5801	6.3090
	GI	10,330	0.4679	0.9219	0.0000	6.8480	0.0000
	ASSETS	10,330	22.5191	1.2091	19.7028	27.6377	22.3449
	AGE	10,330	2.3911	0.5693	0.0000	3.4340	2.3979
	ALR	10,330	3.5800	0.5541	−0.0984	4.5952	3.7052
	R&D	10,330	0.8957	0.8795	0.0000	4.7380	0.8791
Original variables	*GI*	10,330	4.2602	22.5184	1.0000	942.0000	1.0000
	ASSETS	10,330	2.1820	12.1268	1.0000	413.0000	1.0000
	AGE	10,330	3.6481	3.8800	1.0000	114.2047	2.4088
	ALR	10,330	158.0000	433.0000	3.6000	10,100.0000	50.6000
	R&D	10,330	40.7157	17.9592	0.9063	99.0098	40.6594

Note: *ASSETS* original is measured in RMB 100 million (approximately EUR 12.69 million at the People‘s Bank of China central parity rate on August 30, 2024: 1 EUR = CNY 7.8807); the *ROEAVG* variable was winsorized at 1% and 99%.

Source: own

4.2 Correlation analysis of key variables

Tab. 3 shows two key sets of correlations: (1) green innovation (GI) is significantly and positively correlated with firm performance (ROEAVG) (β = 0.0900, $p < 0.01$), which preliminarily supports *H1*; (2) firm performance (ROEAVG) is significantly and positively correlated with firm size (ASSETS) (β = 0.1964, $p < 0.01$), significantly and negatively correlated with asset-liability ratio (ALR) (β = −0.1340, $p < 0.01$) and R&D expenditure (R&D) (β = −0.1212, $p < 0.01$), and has no significant correlation with firm age (AGE) (β = 0.0131, $p > 0.1$).

The results of the VIF test for each variable in Tab. 4 show that the highest VIF value is 1.6200, well below the commonly accepted threshold of 10, suggesting no serious multicollinearity concerns.

Tab. 3: Correlation analysis of key variables

	ROEAVG	*GI*	*ASSETS*	*AGE*	*ALR*	*R&D*
ROEAVG	1.0000					
GI	0.0900***	1.0000				
ASSETS	0.1964***	0.3352***	1.0000			
AGE	0.0131	0.0400***	0.4345***	1.0000		
ALR	−0.1340***	0.2085***	0.4597***	0.2613***	1.0000	
R&D	−0.1212***	0.0614***	0.0824***	0.3090***	0.0155	1.0000

Note: *, **, and *** signify that the results are statistically significant at the 10, 5, and 1%.

Source: own

Tab. 4: Variance inflation factor test for each variable

Variable	VIF	1/VIF
GI	1.1600	0.8630
ASSETS	1.6200	0.6181
AGE	1.4000	0.7140
ALR	1.2900	0.7770
R&D	1.1200	0.8928
Mean VIF	1.3200	–

Source: own

4.3 Multiple linear regression analysis

We estimate the two-way fixed-effects (two-way FE) model (consistent with the definition in Section 3.3: controlling for firm and year FE) to assess the effect of green innovation on firm performance. Tab. 5 provides an overview of the results.

According to Regression (1) in Tab. 5, green innovation positively affects firm performance, even in the absence of controls like firm size, firm age, leverage, and R&D intensity (β = 1.0818, $p < 0.01$), based on this, after adding the control variables, the results still indicate a significant positive relationship between green innovation and firm performance (β = 0.5893, $p < 0.01$), suggesting that green innovation activities contribute positively to firm performance. These findings provide empirical support for *H1*.

It should be pointed out that a relatively low adjusted R^2 (0.1251 in the baseline regression) is quite common in firm-level panel studies focused on innovation. Firm performance is

Tab. 5: Results of regression analyses of main variables

Variable	Regression (1) *ROEAVG*	Regression (2) *ROEAVG*
GI	1.0818*** (0.000)	0.5893*** (0.000)
ASSETS	–	3.0306*** (0.000)
AGE	–	−0.2401 (0.257)
ALR	–	−5.8129*** (0.000)
R&D	–	−1.8030*** (0.000)
Firm	–	Fixed
Year	–	Fixed
_con	5.6676	−39.3494
Adj. R^2	0.0080	0.1251
N	10,330	10,330

Note: *, **, and *** signify that the results are statistically significant at the 10, 5, and 1%; the values in parentheses represent the *p*-value; standard errors are clustered at the firm level to address autocorrelation of residuals within firms.

Source: own

Tab. 6: Results of regression analyses of variables in different regions

Variable	Regression coefficient *ROEAVG*					
	Southern China	Eastern China	Northern China	Southwest China	Northwest China	Northeast China
GI	0.9054*** (0.000)	0.7312*** (0.000)	−0.6839** (0.021)	−0.7559 (0.195)	0.7506 (0.207)	0.1169 (0.909)
ASSETS	3.6077*** (0.000)	3.0894*** (0.000)	2.3791*** (0.000)	3.0548*** (0.000)	3.9440*** (0.000)	2.5329*** (0.000)
AGE	−0.3417 (0.475)	−0.6614** (0.030)	1.3953*** (0.005)	0.5260 (0.557)	−0.3526 (0.758)	−0.9549 (0.388)
ALR	−5.7505*** (0.000)	−5.2277*** (0.000)	5.5292*** (0.000)	−7.2270*** (0.000)	−9.7189*** (0.001)	−8.8984*** (0.000)
R&D	−2.3700*** (0.000)	−1.4643*** (0.000)	−2.2898*** (0.000)	−0.8004 (0.162)	−2.7681*** (0.000)	−2.0380*** (0.004)
Firm	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed
Year	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed
_con	−52.0679	−41.4575	−29.5889	−36.4302	−46.2167	−17.3192
Adj. R^2	0.1376	0.1163	0.1420	0.1178	0.2448	0.2128
N	2,250	5,010	1,670	690	340	370

Note: *, **, and *** signify that the results are statistically significant at the 10, 5, and 1%; standard errors are clustered at the firm level; the values in parentheses represent the *p*-value; according to the Choice Database of Oriental Fortune, China's geographical regions are divided into six regions, namely Southern China, Eastern China, Northern China, Southwest China, Northwest China, and Northeast China.

Source: own

inherently influenced by unmeasured factors (e.g., managerial ability, unexpected market shocks like supply chain disruptions), which are not captured in our model. However, the core independent variable (GI) remains statistically significant, and robustness tests confirm the stability of the green innovation-performance relationship. Thus, the low adjusted R^2 does not undermine the validity of our conclusions.

After controlling for relevant variables, Tab. 6 highlights how firm performance varies in response to green innovation efforts across regions. Among them, enterprise green innovation in Southern China has a positive and significant effect on firm performance ($\beta = 0.9054$, $p < 0.01$), enterprise green innovation in Eastern China also has a positive and significant effect on firm performance ($\beta = 0.7312$, $p < 0.01$), and enterprise green innovation in Northern China has a negative and significant effect on firm performance ($\beta = -0.6839$, $p < 0.05$). None of the enterprise innovations in the Southwest, Northwest, and Northeast regions has a significant contribution to firm performance. *H2* was verified.

4.4 Endogeneity test

Using the lagged value of green innovation as an IV

To examine whether potential endogeneity exists in the relationship between green innovation (GI) and firm performance (ROEAVG), we conduct the Durbin-Wu-Hausman test based on a two-stage least squares (2SLS) framework. The lagged value of green innovation (L.GI) is employed as an instrumental variable, as past innovation decisions are expected to be correlated with current innovation intensity but exogenous to contemporaneous performance shocks.

In the first-stage regression, the instrumental variable (L.GI) shows a significant correlation with the endogenous variable (GI), satisfying the relevance condition. The results of the 2SLS regression (Tab. 7) indicate that GI remains positively and significantly associated with firm performance (β = 0.7525, p = 0.027), suggesting that the core relationship is robust even after correcting for potential endogeneity.

Subsequently, the DWH test is performed to formally verify endogeneity. The test statistic yields F-value equal to 0.3432 with a p-value equal to 0.558, which fails to reject the null hypothesis that the regressors are exogenous. This implies that there is no significant endogeneity bias in the OLS estimation. Therefore, the original OLS model results are consistent and reliable.

Using regional green patent density as an IV

To further examine whether green innovation (GI) suffers from potential endogeneity, the Durbin-Wu-Hausman (DWH) test was conducted using the regional green patent density (GI density) as an instrumental variable (IV). GI density is defined as the ratio of total green patent applications within a province to its corresponding provincial GDP. This variable reflects the regional green innovation environment, regions with higher green patent density are typically characterized by stronger green technology diffusion, policy support, and environmental innovation incentives. Because regional innovation intensity is closely correlated with firms' individual green innovation behavior (satisfying the relevance condition), yet unlikely to directly affect firm performance except through its influence on the firm's own green innovation activities (satisfying the exogeneity condition), GI density is theoretically and empirically a valid instrument.

The instrumental variable estimation results (Tab. 7) indicate that green innovation (GI) remains positively and significantly associated with firm performance (β = 3.5188, p = 0.040).

Tab. 7: Durbin-Wu-Hausman endogeneity test

Variable	*ROEAVG*	*ROEAVG*
IV	Lagged value of green innovation (L.GI)	Regional green patent density (GI density)
GI	0.7525** (0.027)	3.5188** (0.040)
ASSETS	3.7655*** (0.000)	2.7862*** (0.000)
AGE	0.3171 (0.502)	0.8077 (0.129)
ALR	−8.2104*** (0.000)	−7.9262*** (0.000)
R&D	−2.2828*** (0.000)	−2.5980*** (0.000)
Firm	Fixed	Fixed
Year	Fixed	Fixed
_con	−48.6860	−29.7075
N	9,297	10,330
Adj. R^2	0.0648	0.0477
Chi²	–	2.7529 (0.097)
F	0.3432 (0.558)	3.2568 (0.071)

Note: *, **, and *** signify that the results are statistically significant at the 10, 5, and 1%; the values in parentheses represent the p-value.

Source: own

The first-stage regression results confirm the instrument's relevance, with a robust *F*-statistic of 12.2522 (p = 0.001), exceeding the commonly accepted threshold of 10. This suggests that the selected instrument is sufficiently strong to explain variations in GI, thus alleviating concerns of weak instrument bias.

The DWH test further generates a Chi^2 statistic of 2.7529 (p = 0.097) and an *F*-statistic of 3.2568 (p = 0.071). These results indicate marginal endogeneity at the 10% significance level, while no significant bias is detected at the 5% significance level. In other words, while slight endogeneity cannot be entirely excluded, it does not materially distort the causal interpretation. The consistent direction and significance of the coefficient in both OLS and IV estimations confirm that green innovation effectively enhances firm performance, and the observed relationship is not primarily driven by reverse causality or omitted variable bias.

4.5 Robustness test

To address potential regression bias arising from limitations in indicator and model selection, this study performs several robustness tests. These include applying robust and clustered standard errors, incorporating time fixed effects, and replacing key variables. The corresponding findings are reported in Tabs. 8–10.

Tab. 8 shows that using robustness and clustering standard error tests does not change the regression coefficients for green innovation. Although slight variations are observed in standard errors and *t*-values, the *p*-values remain stable and statistically significant.

In addition, the effect of systematic differences between years on the results is

Tab. 8: The test results of robustness and clustering standard error

	GI regression coefficient	Std. error	*t*
Fixed-effect regression	0.5893*** (0.000)	0.1190	4.9500
Robustness criterion error regression	0.5893*** (0.000)	0.1171	5.0300
Standard error regression for province clustering	0.5893** (0.013)	0.2236	2.6400
Standard error regression for year clustering	0.5893** (0.013)	0.1201	4.9000

Note: *, **, and *** signify that the results are statistically significant at the 10, 5, and 1%; the values in parentheses represent the *p*-value.

Source: own

Tab. 9: The regression results of controlling for time effects

	GI regression coefficient	Std. error	*t*
Fixed-effect regression	0.5893*** (0.000)	0.1190	4.9500
Controlling for time effect regression	0.5668** (0.011)	0.2226	2.5500

Note: *, **, and *** signify that the results are statistically significant at the 10, 5, and 1%; the values in parentheses represent the *p*-value.

Source: own

eliminated by controlling for time (year) fixed effects. The regression results in Tab. 9 show that the *p*-value remains significant after controlling for time effects.

Finally, to further test the robustness of the model, a regression was conducted using return on invested capital (ROIC) as an alternative to average return on equity (ROEAVG) and also employed a GI with a lag of three periods. The robustness check in Tab. 10 confirms that significance persists even with alternative variable specifications and a lag of three periods, lending further credibility to the analysis.

Tab. 10: The regression results for the substitute variable and three-period lagged GI

Variable	Regression coefficient		
	ROEAVG (1)	*ROEAVG* (2)	ROIC
GI	0.5893*** (0.000)	0.8986*** (0.000)	0.3460*** (0.000)
ASSETS	3.0306*** (0.000)	6.3509*** (0.000)	2.2752*** (0.000)
AGE	−0.2401 (0.257)	−3.5734*** (0.004)	0.0457 (0.796)
ALR	−5.8129*** (0.000)	−9.3978*** (0.000)	−4.3421*** (0.000)
R&D	−1.8030*** (0.000)	−2.1092*** (0.000)	−1.5477*** (0.000)
_con	−39.3494	−91.9805	−28.6573
N	10,330	10,330	10,330
Adj. R^2	0.1251	0.1422	0.1067

Note: *, **, and *** signify that the results are statistically significant at the 10, 5, and 1%; the values in parentheses represent the *p*-value; *ROEAVG* (2) using three-period lagged GI (t − 3).

Source: own

5 Discussion

5.1 Enterprise green innovation has a significant positive effect on firm performance

The empirical results show that there is a significant positive relationship between green innovation and firm performance, implying that with the increase of green innovation activities of enterprises, firm performance improves significantly. This result supports the current focus on green innovation in the field of business management and suggests that in the modern market, enterprises can improve market competitiveness and economic efficiency to a certain extent through green innovation (Padilla-Lozano & Collazzo, 2022). Green innovation not only helps reduce energy consumption and environmental pollution, but also may enhance the brand image and market reputation of the enterprise, potentially resulting in increased sales and greater market share (Chen et al., 2023; Novitasari & Tarigan, 2022). In addition, green innovation can promote technological advancement and process optimization within the enterprise, increasing productivity and thus reducing operating costs, which can translate into improved financial performance (Yang et al., 2022).

In many countries and regions, governments promote the implementation of green innovation by enterprises through environmental regulations, tax incentives, and subsidy policies (Jiang et al., 2023; C. Wang et al., 2022). These empirical findings highlight the effectiveness of such policies in encouraging enterprises to invest in green technologies and projects. Government support not only helps enterprises reduce the pressure of upfront investment in green innovation but also promotes long-term sustainable development of enterprises by building a favorable policy environment.

Increasing consumer demand for environmentally sustainable products compels firms to adopt green innovations in the market with green innovations to meet consumer expectations. The high β value in the empirical results reflects this trend, suggesting that enterprises' green innovation efforts can be directly translated into competitive market advantages and better profitability performance.

The performance and outcomes of green innovation may vary among enterprises in different industries. For example, green innovation in high-pollution sectors such as manufacturing, energy, and chemicals may lead to more significant performance improvements, as these industries tend to face stricter environmental regulations and greater market pressure (Gao et al., 2023). For service and high-tech industries, on the other hand, green innovation may be more in the form of product and service differentiation, which can enhance firm performance enhancing customer loyalty and strengthening brand equity (Borah et al., 2023).

5.2 Disparities in green innovation outcomes across regional business contexts

Regional heterogeneity in the impact of green innovation on firm performance is shaped by multiple factors, including the level of economic development, industrial composition, policy incentives, and technological capabilities. In economically developed regions characterized by strong market demand and robust policy support, green innovation tends to contribute more significantly to firm performance. Conversely, in relatively underdeveloped regions with traditional industrial structures, the effects of green innovation are often less pronounced or may even have negative outcomes.

The South and East China regions represent the most economically advanced areas in the country, where firms possess greater resources and stronger capacities for green innovation. As the pace of economic growth accelerates in these regions, enterprises exhibit higher competitiveness in green innovation investment, stronger capabilities in absorbing and applying new technologies, and, consequently, greater improvements in firm performance. Given that green innovation often entails substantial financial investment and technological accumulation, it is more feasible in economically developed regions (Li & Liu, 2024). Moreover, these areas demonstrate higher market demand for green products and services, and local governments provide more vigorous policy support for green development.

Specifically, the Chinese government has introduced a range of targeted policies aligned with the developmental characteristics of East and South China. In the domain of green finance, the State Administration of Foreign Exchange launched pilot programs for green external debt in 16 provinces and municipalities, including Shanghai, Guangdong, and Zhejiang. Leveraging this policy, Qingdao promoted the issuance of RMB 380 million (approximately EUR 45.83 million at the People's Bank of China central parity rate on August 28, 2025: 1 EUR = CNY 8.2915) in low-carbon loans by financial institutions, while Ningbo facilitated green external debt projects supporting lithium battery manufacturing in the new energy sector. In terms of industrial upgrading, the Yangtze River Delta and the Guangdong-Hong Kong-Macao Greater Bay Area have been designated as key zones for green development, establishing cross-regional emission trading mechanisms to guide industries such as electronics and automobiles toward resource efficiency through "green supply chains". Meanwhile, stricter environmental regulations, such as special emission limits for steel and chemical industries in the Yangtze River Delta, have compelled firms to pursue cleaner production transformations. These policies, when combined with regional economic and industrial advantages, form a synergistic framework that provides enterprises with financial, technological, and market incentives for green innovation. As a result, firms in East and South China have achieved greater performance gains from green innovation initiatives.

The industrial structures of South and East China are dominated by manufacturing and high-tech sectors, which are inherently innovative and environmentally sensitive. Green innovation enhances production efficiency, reduces resource consumption and pollution emissions, and consequently improves firm performance. Firms in these regions lead in technological development and application, enabling them to convert green innovation into productivity more effectively. In contrast, in the Southwest, Northwest, and Northeast regions-where traditional industries remain dominant and technological development lags behind, the input-output efficiency of green innovation is relatively low, leading to an insignificant impact on firm performance (Liu et al., 2021; S. Wang et al., 2021).

To address these regional disparities, the government has implemented preferential policies to compensate for insufficient green innovation resources. In the Southwest and Northwest, based on the Catalogue of Encouraged Industries in the Western Region (2025 Edition) (NDRC, 2024), industries such as green coal chemical production and photovoltaic module manufacturing have been included in the encouraged category, receiving tax and land incentives. For instance, Guizhou is advancing an integrated "phosphate-power-chemical" circular economy, while Sichuan is realizing ecological value transformation through the Giant Panda National Park carbon sink project. Additionally, the "West-to-East Power Transmission" initiative has expanded wind and solar power capacity in Northwest China and, through the establishment of "enclave industrial parks", facilitated the transfer of green technologies from eastern to central and western regions.

In the Northeast, the government has promoted the green transformation of traditional industries through the issuance of special-purpose bonds. For example, Anshan Iron and Steel in Liaoning invested in a hydrogen-based shaft furnace project, while Jilin has been driving the transformation of its automobile industry toward new energy vehicles. A "green transformation fund" has also been established to attract private capital participation in state-owned enterprise reform. However, due to long industrial transformation cycles and weak technological foundations, the R&D intensity of green technologies among large-scale industrial enterprises in the Northeast remains below the national average, and the positive effects of green innovation on firm performance have not yet been fully realized.

Furthermore, some regions with abundant natural resources and lower environmental pressures exhibit weaker motivation for green innovation or less tangible results (Shuhong Wang et al., 2022). For example, while the Southwest and Northwest are rich in ecological and energy resources, historically lax environmental constraints have reduced firms' incentives for green innovation. At present, the government is gradually transforming these resource advantages into drivers of green innovation through ecological compensation mechanisms, such as the horizontal ecological protection compensation system in the Yangtze River Basin, and resource recycling initiatives, including Heilongjiang's "straw-to-meat" program. Nevertheless, in the short term, these regions still lag behind East and South China, which further explains why the contribution of green innovation to firm performance in the Southwest, Northwest, and Northeast remains limited.

Conclusions

This study examined the impact of green innovation on firm performance using panel data from 1,033 manufacturing firms listed on the Shanghai and Shenzhen stock exchanges between 2014 and 2023. The empirical findings confirm a statistically significant positive relationship between green innovation and firm performance, though the magnitude and direction of this effect vary across regions.

In developed areas such as Southern and Eastern China, green innovation positively contributes to firm performance, reflecting the presence of supportive institutional environments, advanced industrial infrastructure, and higher market receptiveness to green products. Conversely, in Northern China, the relationship is negative likely due to high implementation costs and regulatory pressures outweighing short-term benefits. In less-developed regions like the Southwest, Northwest, and Northeast, the effect is insignificant, suggesting an underdeveloped green innovation ecosystem.

Theoretical contributions. This research refines the RBV framework by viewing green innovation as a key intangible asset that shapes performance outcomes under varying regional circumstances. It enriches empirical understanding of how institutional environments moderate the effectiveness of green innovation in emerging markets, especially within the Chinese manufacturing sector. It offers a multilevel, regionally differentiated perspective often overlooked in prior studies, thus addressing key gaps in the extant literature.

Practical implications. Managers should align green innovation strategies with regional capabilities. In regions with mature infrastructure and supportive policies, proactively investing in substantive (high-quality) green innovation helps build competitive advantage.

In underdeveloped regions, prioritizing capability-consistent, incremental pathways and actively seeking targeted public programs (e.g., green credit, tax offsets, demonstration projects) is advisable. Policymakers should favor predictable enforcement and disclosure regimes, which are associated with more numerous and more influential green patents and better market pricing of eco-innovation. Providing adoption cost-reducing incentives tied to substantive innovation (to curb greenwashing), building regional green-innovation platforms, and enhancing knowledge spillovers can help narrow disparities in the transition toward sustainability.

Finally, we delineate the study's boundaries and external validity to guide interpretation and future work. This study's inference is bounded by its empirical domain and measurement choices. First, the sample focuses on A-share listed manufacturers in China; hence, results may not extrapolate to SMEs, unlisted and private firms, or service intensive industries where intangible, non-patent eco-innovations are more prevalent. Caution is warranted when generalizing these results: SMEs have weaker R&D capacity, unlisted firms may face stronger financing constraints, and service industries often engage in non-patented green innovations, factors that may alter the observed green innovation, performance relationship. Second, our green innovation metrics are primarily patent-based (applications, grants, forward citations). While standard in the literature, they tilt toward technological outputs and may underrepresent organizational or process practices that are not patented. Citation-window truncation and self-citation exclusion, although addressed in robustness checks, can still leave residual measurement error. Third, the regulatory and capital market context of China during the sample period conditions both firms' incentives and investors' pricing of green innovation; external validity to jurisdictions with weaker enforcement or different disclosure regimes should be made cautiously. Finally, although we mitigate simultaneity using lag structures, rich controls, and fixed effects, we refrain from strong causal language and interpret coefficients as average partial effects under stated assumptions. Future research could integrate complementary data sources, such as environmental management certifications, survey-based indicators, or firm-level ESG disclosures, to provide a more comprehensive understanding of non-technological green innovation. Moreover, it could triangulate patent-based indicators with trademarks, environmental management certifications, and survey based measures, and extend to multi-country panels to probe boundary conditions more fully.

Acknowledgments: *This study was supported by grant No. LRB202408 Optimal Contest Design: Sabotage-Proof and Discrimination Policies of the Doctoral Research Start-Up Foundation of Liming Vocational University and grant No. 2022H18 Research on the Impact of Entrepreneurship on the Innovation Performance of Private Enterprises of the Quanzhou Social Science Planning Project.*

References

Abadzhiev, A., Sukhov, A., Sihvonen, A., & Johnson, M. (2022). Managing the complexity of green innovation. *European Journal of Innovation Management*, *25*(6), 850–866. https://doi.org/10.1108/EJIM-02-2022-0098

Bai, Y., Song, S., Jiao, J., & Yang, R. (2019). The impacts of government R&D subsidies on green innovation: Evidence from Chinese energy-intensive firms. *Journal of Cleaner Production*, *233*, 819–829. https://doi.org/10.1016/j.jclepro.2019.06.107

Baltagi, B. H. (2021). *Econometric analysis of panel data*. Springer Nature. https://doi.org/10.1007/978-3-030-53953-5

Barney, J. B., Ketchen, D. J., & Wright, M. (2021). Resource-based theory and the value creation framework. *Journal of Management*, *47*(7), 1936–1955. https://doi.org/10.1177/01492063211021655

Block, J., Lambrecht, D., Willeke, T., Cucculelli, M., & Meloni, D. (2025). Green patents and green trademarks as indicators of green innovation. *Research Policy*, *54*(1), 105138. https://doi.org/10.1016/j.respol.2024.105138

Borah, P. S., Dogbe, C. S. K., Pomegbe, W. W. K., Bamfo, B. A., & Hornuvo, L. K. (2023). Green market orientation, green innovation capability, green knowledge acquisition and green brand positioning as determinants of new product success. *European Journal of Innovation Management*, *26*(2), 364–385. https://doi.org/10.1108/EJIM-09-2020-0345

Chen, Z., Hao, X., & Chen, F. (2023). Green innovation and enterprise reputation value. *Business Strategy and the Environment*, *32*(4), 1698–1718. https://doi.org/10.1002/bse.3213

Cheng, Q., Lin, A.-P., & Yang, M. (2025). Green innovation and firms' financial and environmental performance: The roles of pollution prevention versus control. *Journal of Accounting and Economics*, *79*(1), 101706. https://doi.org/10.1016/j.jacceco.2024.101706

De Oliveira, N. A., & Basso, L. F. C. (2024). The impact of value creation (Tobin's Q), total shareholder return (TSR), and survival (Altman's Z) on credit ratings. *International Journal of Financial Studies*, *12*(2), 44. https://doi.org/10.3390/ijfs12020044

Do, T. K. (2024). Asset redeployability and green innovation. *Journal of Financial Stability*, *72*, 101270. https://doi.org/10.1016/j.jfs.2024.101270

Favot, M., Vesnic, L., Priore, R., Bincoletto, A., & Morea, F. (2023). Green patents and green codes: How different methodologies lead to different results. *Resources, Conservation & Recycling Advances*, *18*, 200132. https://doi.org/10.1016/j.rcradv.2023.200132

Gao, J., Feng, Q., Guan, T., & Zhang, W. (2023). Unlocking paths for transforming green technological innovation in manufacturing industries. *Journal of Innovation & Knowledge*, *8*(3), 100394. https://doi.org/10.1016/j.jik.2023.100394

He, Q., & Qiu, B. (2025). Environmental enforcement actions and corporate green innovation. *Journal of Corporate Finance*, *91*, 102711. https://doi.org/10.1016/j.jcorpfin.2024.102711

Huang, Z., & Xiao, Z. (2023). Dynamic capabilities, environmental management capabilities, stakeholder pressure and eco-innovation of Chinese manufacturing firms: A moderated mediation model. *Sustainability*, *15*(9), 7571. https://doi.org/10.3390/su15097571

Jiang, Z., Xu, C., & Zhou, J. (2023). Government environmental protection subsidies, environmental tax collection, and green innovation: Evidence from listed enterprises in China. *Environmental Science and Pollution Research*, *30*(2), 4627–4641. https://doi.org/10.1007/s11356-022-22538-3

Kemp, R., Rammer, C., & Arundel, A. (2023). Measuring environmental (eco-) innovation. In *Handbook of innovation indicators and measurement* (pp. 177–196). Edward Elgar Publishing. https://doi.org/10.4337/9781800883024.00020

Levchenko, I., Losonczi, P., Britchenko, I., Vazov, R., Zaiats, O., Volodavchyk, V., Humeniuk, I., & Shumilo, O. (2021). Development of a method for targeted financing of economy sectors through capital investment in the innovative development. *Eastern-European Journal of Enterprise Technologies*, *5*(13(113)), 6–13. https://doi.org/10.15587/1729-4061.2021.243235

Li, X., & Liu, X. (2024). Explaining the changes in the green technology innovation efficiency of construction enterprises. *Humanities and Social Sciences Communications*, *11*(1), 1–15. https://doi.org/10.1057/s41599-024-03966-7

Liu, J., Long, F., Chen, L., Li, L., Zheng, L., & Mi, Z. (2025). Exploratory or exploitative green innovation? The role of different green fiscal policies in motivating innovation. *Technovation*, *143*, 103207. https://doi.org/10.1016/j.technovation.2025.103207

Liu, K., Qiao, Y., & Zhou, Q. (2021). Analysis of china's industrial green development efficiency and driving factors: Research based on MGWR. *International Journal of Environmental Research and Public Health*, *18*(8), 3960. https://doi.org/10.3390/ijerph18083960

Liu, L., Feng, A., & Liu, M. (2024). The effect of green innovation on corporate financial performance: Does quality matter? *Finance Research Letters*, *62*, 105255. https://doi.org/10.1016/j.frl.2024.105255

Liu, M., Liu, L., & Feng, A. (2024). The impact of green innovation on corporate performance: An analysis based on substantive and strategic green innovations. *Sustainability*, *16*(6), 2588. https://doi.org/10.3390/su16062588

Maihaiti, G., Hunjra, A. I., Alharbi, S. S., & Zhao, S. (2025). Climate-resilient cities for green innovation in enterprises: New perspective based on Fintech. *International Review of Economics & Finance*, *101*, 104251. https://doi.org/10.1016/j.iref.2025.104251

Meng, F., Xu, Y., & Zhao, G. (2020). Environmental regulations, green innovation and intelligent upgrading of manufacturing enterprises: Evidence from China. *Scientific Reports*, *10*(1), 14485. https://doi.org/10.1038/s41598-020-71423-x

Mirza, S. S., Xu, N., Corbet, S., & Scrimgeour, F. (2025). Green innovation under financial and policy uncertainty: Evidence from China. *Research in International Business and Finance*, *76*, 102856. https://doi.org/10.1016/j.ribaf.2025.102856

National Development and Reform Commission (NDRC). (2024). *Catalogue of Encouraged Industries in the Western Region (2025 Edition) (Order No. 28 of 2024)*. People's

Republic of China. https://www.ndrc.gov.cn/xxgk/zcfb/fzggwl/202411/t20241129_1394744.html

Novitasari, M., & Tarigan, Z. J. H. (2022). The role of green innovation in the effect of corporate social responsibility on firm performance. *Economies*, 10(5), 117. https://doi.org/10.3390/economies10050117

Ostapenko, I., Zadykhaylo, D., Lyseiuk, A., Tunitska, Y., & Sierova, L. (2024). Global practices and experiences in developing a green economy amid financial crises. *Grassroots Journal of Natural Resources*, *7*(3), 244–270. https://doi.org/10.33002/nr2581.6853.070314

Padilla-Lozano, C. P., & Collazzo, P. (2022). Corporate social responsibility, green innovation and competitiveness – Causality in manufacturing. *Competitiveness Review: An International Business Journal*, *32*(7), 21–39. https://doi.org/10.1108/CR-12-2020-0160

Sadiq, M., & Gebba, T. R. A. (2022). Financial performance, firm value, transparency and corporate governance. Evidences from family-owned business in UAE. *Transnational Corporations Review*, *14*(3), 286–296. https://doi.org/10.1080/19186444.2021.1938496

Singh, S. K., del Giudice, M., Chierici, R., & Graziano, D. (2020). Green innovation and environmental performance: The role of green transformational leadership and green human resource management. *Technological Forecasting and Social Change*, *150*, 119762. https://doi.org/10.1016/j.techfore.2019.119762

Tian, J., Dong, Y., Vagnani, G., & Liu, P. (2023). Green innovation and the stock market value of heavily polluting firms: The role of environmental compliance costs and technological collaboration. *Business Strategy and the Environment*, *32*(7), 4938–4953. https://doi.org/10.1002/bse.3401

Valaskova, K., Gajdosikova, D., & Pavic Kramaric, T. (2022). How important is the business environment for the performance of enterprises? Case study of selected European countries. *Central European Business Review*, *11*(4), 85–110. https://doi.org/10.18267/j.cebr.300

Vimalnath, P., Tietze, F., Jain, A., Gurtoo, A., Eppinger, E., & Elsen, M. (2022). Intellectual property strategies for green innovations – An analysis of the European Inventor Awards. *Journal of Cleaner Production*, *377*, 134325. https://doi.org/10.1016/j.jclepro.2022.134325

Wang, B., Xu, L., & Chen, L. (2021). Factors affecting green innovation: An analysis of patent and regional heterogeneity. *Chinese Journal of Population, Resources and Environment*, *19*(1), 12–21. https://doi.org/10.1016/j.cjpre.2021.12.002

Wang, C., Chen, P., Hao, Y., & Dagestani, A. A. (2022). Tax incentives and green innovation – The mediating role of financing constraints and the moderating role of subsidies. *Frontiers in Environmental Science*, *10*, 1067534. https://doi.org/10.3389/fenvs.2022.1067534

Wang, H., Yang, J., & Zhu, N. (2024). Does tax incentives matter to enterprises' green technology innovation? The mediating role on R&D investment. *Sustainability*, *16*(14), 5902. https://doi.org/10.3390/su16145902

Wang, Shouwen, Abbas, J., Sial, M. S., Álvarez-Otero, S., & Cioca, L.-I. (2022). Achieving green innovation and sustainable development goals through green knowledge management: Moderating role of organizational green culture. *Journal of Innovation & Knowledge*, *7*(4), 100272. https://doi.org/10.1016/j.jik.2022.100272

Wang, Shuhong, Wang, X., & Lu, B. (2022). Is resource abundance a curse for green economic growth? Evidence from developing countries. *Resources Policy*, *75*, 102533. https://doi.org/10.1016/j.resourpol.2021.102533

Wang, S., Zhang, Y., & Wen, H. (2021). Comprehensive measurement and regional imbalance of China's green development performance. *Sustainability*, *13*(3), 1409. https://doi.org/10.3390/su13031409

Yang, H., Li, L., & Liu, Y. (2022). The effect of manufacturing intelligence on green innovation performance in China. *Technological Forecasting and Social Change*, *178*, 121569. https://doi.org/10.1016/j.techfore.2022.121569

Zhang, Z., Duan, H., Shan, S., Liu, Q., & Geng, W. (2022). The impact of green credit on the green innovation level of heavy-polluting enterprises – Evidence from China. *International Journal of Environmental Research and Public Health*, *19*(2), 650. https://doi.org/10.3390/ijerph19020650

Zheng, L., & Iatridis, K. (2022). Friends or foes? A systematic literature review and meta-analysis of the relationship between eco-innovation and firm performance. *Business Strategy and the Environment*, *31*(4), 1838–1855. https://doi.org/10.1002/bse.2986

Strategic intangibles in small family firms: The mediating role of innovativeness in resource-constrained environments

Syed Amjad Hussain[1], Salman Iqbal[2], Vladimir Bures[3], Abdul Waheed[4], Alzbeta Docekalova[5], Marek Zanker[6]

[1] University of Central Punjab, Faculty of Management Sciences, Department of Management & Entrepreneurship, Pakistan, ORCID: 0000-0002-6719-0078, amjadhussain@ucp.edu.pk;
[2] University of Central Punjab, Faculty of Management Sciences, Department of Management & Entrepreneurship, Pakistan, ORCID: 0000-0001-6595-9288, salman.iqbal@ucp.edu.pk;
[3] University of Hradec Králové, Faculty of Informatics and Management, Department of Information Technologies, Czech Republic, ORCID: 0000-0001-7788-7445, vladimir.bures@uhk.cz;
[4] University of Central Punjab, Faculty of Management Sciences, Department of Management & Entrepreneurship, Pakistan, ORCID: 0000-0003-3199-1014, a.waheed@ucp.edu.pk;
[5] University of Hradec Králové, Faculty of Informatics and Management, Department of Information Technologies, Czech Republic, ORCID: 0009-0007-3602-2841, alzbeta.docekalova@uhk.cz;
[6] University of Hradec Králové, Faculty of Informatics and Management, Department of Information Technologies, Czech Republic, ORCID: 0000-0002-2745-4868, marek.zanker@uhk.cz (corresponding author).

Abstract: *This study addresses a critical gap in the literature by examining how family-based brand identity and entrepreneurial orientation (specifically, proactiveness and risk taking) influence the performance of small family firms operating in emerging markets. Grounded in the resource-based view (RBV), these constructs are conceptualized as strategic intangible resources that underpin competitive advantage. A cross-sectional survey of 460 privately held small family firms in Pakistan was conducted, using validated scales for family-based brand identity, entrepreneurial orientation (EO) dimensions, innovativeness, and performance. Following a pilot test, the data underwent confirmatory factor analysis and structural equation modeling (SEM) in AMOS 24.0 to assess direct and mediating relationships. The findings reveal that cultivating a family-based brand identity and entrepreneurial orientation significantly contributes to improved firm performance by recognizing the family as the brand and EO as family firm resources via the RBV, by assessing their impact on performance, and by mediating the effect of innovativeness. The study extends RBV theory through the identification of family-based brand identity and EO as firm-specific resources in small family firms and demonstrates innovativeness as a mediator.*

Keywords: *Family branding, entrepreneurial orientation, risk-taking, proactiveness, innovativeness, resource based view.*

JEL Classification: *L26, M10, D22, L25, O53, G32.*

APA Style Citation: Hussain, S. A., Iqbal, S., Bures, V., Waheed, A., Docekalova, A. & Zanker, M. (2026). Strategic intangibles in small family firms: The mediating role of innovativeness in resource-constrained environments. *E&M Economics and Management, 29*(2), pp. 124–140 https://doi.org/10.15240/tul/001/2026-2-009

Introduction

Small family firms (SFFs) are a dominant force in the global economy, making up a significant share of businesses worldwide. They contribute approximately 70%–90% of global GDP and are responsible for 50%–80% of job creation in most countries. Their impact is especially notable in both developed and emerging markets (Luo, 2019). These firms have gained recognition for their outstanding performance, particularly in challenging economic circumstances (Buchanan et al., 2023; Fang et al., 2025). They demonstrate competitive advantage in areas such as product and service excellence, innovation, and financial accomplishments (Smith et al., 2024). On a global scale, family firms are perceived to be less susceptible to failure compared to nonfamily firms (Madanoglu et al., 2020). One of the reasons for their success is their ability to establish themselves as a brand in the business community (Craig et al., 2008). Furthermore, SFFs are often seen as more trustworthy, responsible, and customer-oriented than nonfamily firms (Beck & Prügl, 2018).

Despite its intuitive appeal, empirical research has not yet tested whether small family firms leverage a family brand identity to secure the performance advantages documented in larger family firms. Similarly, small family firms, being smaller in size than their listed counterparts, face significant market challenges such as limited access to capital, narrower business networks, and weaker economies of scale. This financial conservatism significantly limits marketing expenditures, compelling owners to devise low-cost, creative promotional strategies for which they often lack the requisite expertise and institutional backing (Dabić et al., 2023). Concurrently, the scope of competition from larger domestic and multinational firms erodes market share, as these entities leverage superior resource pools and scale efficiencies to outcompete smaller family-owned operations (Luo, 2019). The deficiency of robust business networks further hampers these firms' access to essential information and political or industry connections, deepening their marginalization in transitional economies (Luo, 2019).

The challenge of adopting and implementing new technologies represents a critical barrier to the competitiveness of family firms, especially under crisis conditions when agility and innovation are paramount (Dabić et al., 2023). Restricted access to capital not only limits its direct investment in digital tools and machinery but also reduces opportunities to engage in collaborative research and development networks that facilitate technological diffusion (Luo, 2019). The absence of specialized managerial capabilities and formal governance structures compounds this issue, leading to delayed or underinvested technology initiatives that widen the performance gap with more advanced competitors.

This research aims to fill this gap by addressing the following question. How do family-based brand identity and entrepreneurial orientation (specifically proactiveness and risk taking) affect the performance of small family firms in Pakistan? To clarify this underexplored context, non-listed small family firms in Pakistan, an environment marked by limited formal financing and networks, were surveyed to evaluate how these resources drive innovativeness and performance under the RBV framework.

1 Theoretical background

1.1 Branding and entrepreneurial orientation as resources

The RBV systematically organizes intangible assets, such as branding, entrepreneurial orientation (EO), organizational culture, market knowledge, and human capital, as strategic resources that underpin firm performance (Huybrechts et al., 2011) and competitive advantage. Within this framework, intangible resources are particularly salient for family firms and small and medium-sized enterprises (SMEs) (Sarsiti & Minarni, 2024) owing to their limited tangible endowments and reliance on nonphysical assets for differentiation. Branding and EO emerge as pivotal components of this asset portfolio, linking firm-specific capabilities to market positioning and growth trajectories. The integration of brand-related symbols and strategic mindedness contributes to sustained value creation, especially under competitive and resource-constrained conditions.

Brand orientation positions the brand as a core strategic resource by embedding intangible values and symbolic associations into the firm's overarching strategy, thereby fostering strong brand equity (Wiid et al., 2025). Empirical studies demonstrate that brand reputation is further reinforced by a firm's entrepreneurial orientation, with the innovativeness dimension playing a particularly crucial role in elevating reputational capital (Rua et al., 2018).

Entrepreneurial orientation itself constitutes an essential intangible resource, characterized by innovativeness, risk taking, and proactiveness, directly driving firm performance, growth, and competitive advantage (Rua et al., 2018). Beyond its direct effects, EO facilitates the acquisition and effective exploitation of other intangible assets (such as market knowledge, relational networks, and brand reputation), thereby amplifying its impact in dynamic and international markets (Baquero, 2024). In environments characterized by rapid change and heightened uncertainty, EO enables firms to identify and capitalize on emerging opportunities, reinforcing their strategic adaptability (Lee & Chu, 2013). The multidimensional nature of EO thus interacts with existing resources, fostering a capability-building process that underpins sustained competitive advantage (Buttar & Kocak, 2011). This perspective underscores the dual role of EO as both a direct performance driver and an enabler of broader resource configurations.

The crux of the matter lies in acknowledging the pivotal importance of intangible resources within family firms (Huybrechts et al., 2011). In this context, comprehending the intricacies of resource management, leveraging, bundling, and their seamless integration (Hussain, 2017) assumes paramount significance. In the realm of improving the effectiveness of family firms, it is crucial to examine family branding and entrepreneurial orientation (proactiveness and risk taking) as key drivers and significant organizational resources (Arzubiaga et al., 2019). These inquiries revolve around characterizing these qualities as resources, understanding the conditions under which they generate value, and examining the resulting capabilities and returns for the firm.

The research systematically tackles these questions, clarifying how these identified resources impact the performance of SFFs. Amid the broader scope of firm resources, which includes aspects such as location, access to capital, competencies, and culture, the study categorizes family branding and entrepreneurial orientation (proactiveness and risk taking) as intangible resources falling under the categories of “organizational capital resources” and “process capital resources”, respectively. This classification arises from their significant relevance to the family’s reputation within the community, establishing the business as a ‘family’ brand, and entrepreneurial orientation as part of the firm’s culture and practices. The inherent value of a resource remains unrealized until its potential for value enhancement is understood. Consequently, the influence of these resources on firm performance is empirically assessed, illuminating their transformative capacity.

1.2 Hypotheses development

At a fundamental level, a company’s brand identity, representing what the company stands for, stands out as a crucial intangible asset and often serves as a key driver of competitive advantage (Melewar et al., 2014). Entrepreneurial orientation refers to the company’s strategies and activities geared toward actively pursuing new business opportunities (Rauch et al., 2009), such as offering new products or exploring new markets.

Family branding

Despite support for ongoing initiatives, prior research has offered limited insights into how family firms strategically utilize their identity linked to the family name. The scarcity of academic exploration in mainstream literature has impeded a comprehensive understanding of family brand identification. In the current study, the significance of family brand identification is underscored for the success of entrepreneurial ventures and small to medium-sized firms. Aligned with the RBV framework, family brand identity is posited as a rare, valuable, imperfectly imitable, and non-substitutable resource capable of conferring a competitive advantage. Utilizing the RBV-informed unified systems perspective of family firm performance (Barbera et al., 2023), family-based brand identity emerges as a contributing resource and capability, offering SFFs a potential edge by emphasizing their distinct familial character through family branding. This leads us to formulate the following hypothesis:

H1: Family branding as an organizational capital resource is positively related to the performance of small family firms.

Entrepreneurial orientation

Entrepreneurial orientation comprises two key aspects: proactiveness and risk taking (Hult et al., 2004). Entrepreneurial orientation refers to a company's strategies and activities geared toward actively pursuing new opportunities (Rauch et al., 2009). Acting proactively, characteristic of risk-taking organizations, has the potential to disrupt traditional leadership and structures, opening new business avenues and positioning SFFs ahead of competitors through proactive environmental scanning.

Proactiveness is the first element of EO and pertains to a firm's ability to foresee future challenges, changes, and demands. It represents an organizational initiative aimed at seizing opportunities, thereby allowing companies to stay ahead of their competitors through the introduction of innovative products and services (Rauch et al., 2009).

Within family firms, these components represent the intangible benefits and non-monetary advantages derived from ownership of a particular firm. The second dimension of entrepreneurial orientation, risk taking, operates on the premise that entrepreneurial actions are often accompanied by uncertainty (Guo & Jiang, 2020). Embracing uncertainty is inherent in entrepreneurial ventures, where entrepreneurs willingly accept the unknown for the sake of novel undertakings. Organizational behavior, considered a resource, uniquely contributes to a firm's ability to positively engage in controlled and calculated risk ventures (Morris et al., 2010). A firm's proclivity for risk taking fosters learning, enhancing entrepreneurs' capabilities to navigate uncertain situations and ultimately promoting entrepreneurial behavior within the organization. This trait is deemed valuable in dynamic and ambiguous environments, positively influencing a firm's reputation and performance (Kreiser & Davis, 2010). Following Covin and Wales (2012), proactiveness and risk taking are identified as process-oriented dimensions of EO to avoid conceptual overlap with innovativeness, which is examined here as a distinct capability mediating performance.

H2: Small family firms' proactiveness as a process capital resource is positively related to the performance of small family firms.

H3: Small family firms' risk taking as a process capital resource is positively related to the performance of small family firms.

Innovativeness

The concept of innovativeness involves an organization engaging in experimental behaviors to create new products or services through research and development efforts (Rauch et al., 2009). Research indicates that increased investment in innovation can lead to a sustainable competitive advantage (Chatzoglou & Chatzoudes, 2018). SFF investing in innovation are found to have high potential for performance and growth. Innovative capacity helps these SFFs strengthen their market position over time by facilitating adaptation and the initiation of changes in their markets and industry. Therefore, innovative capacity emerges as a crucial resource contributing to family firm success by promoting entrepreneurial activities that enhance distinctiveness and profitability.

Innovativeness as mediator

Family-based brand identity enhances employee identification and external trust, thereby motivating idea sharing and investment in novel processes. Proactiveness facilitates systematic environmental scanning and early experimentation, while risk taking lowers internal barriers to the adoption of new ideas, collectively reinforcing organizational innovativeness.

Innovativeness stands out as a vital factor determining firm performance (Kyrgidou & Spyropoulou, 2013). Firms possessing the ability to innovate can develop a competitive advantage, leading to positive outcomes (Chatzoglou & Chatzoudes, 2018). Innovativeness also serves as a strategic approach for firms to navigate changes in both internal and external environments. To effectively respond to a turbulent environment, nurturing innovativeness becomes crucial, playing a critical role in gaining a competitive edge and achieving superior performance (Hult et al., 2004).

A family-based brand identity can impact performance through various competitive mechanisms, though the precise nature of this effect is not entirely clear. This study concentrates on innovativeness, as it signifies the competitive advantage sought through the behavior, culture, or tendencies of the firm (Mishra, 2021).

Hult et al. (2004) define innovativeness as an organization's ability to introduce new processes, products, or ideas. Organizational innovativeness is considered an aspect of organizational culture (Riivari & Lämsä, 2014). A comprehensive approach defines innovativeness as a firm's inclination to embrace and support new ideas, novelty, experimentation, and creative processes that may lead to new products, services, or technological advancements

(Hügel, 2019). Given the increasing complexity of managerial work in modern technological environments (Bureš & Čech, 2007), small firms rely significantly on their entrepreneurial orientation, characterized by proactiveness and risk taking, to gain a competitive advantage (Sunday Aneke & Ja'afaru Garba, 2024). Various factors directly influence innovativeness, with organizational culture being one of the key antecedents. Organizational culture has the potential to foster a receptivity to new ideas and innovation within an organization (Valentin & Călin, 2024). Therefore, innovativeness, described as the action-based capacity to introduce and implement creative ideas within a firm, appears closely linked to entrepreneurial orientation. These theoretical perspectives and empirical findings suggest the mediating role of innovativeness. Consequently, the following hypotheses are proposed:

H4: Innovativeness mediates the relationship between family firm branding and performance in small family firms.

H5: Innovativeness mediates the relationship between proactiveness and performance in small family firms.

H6: Innovativeness mediates the relationship between risk taking and performance in small family firms.

The specific functions of organizational capital resources (family branding), process capital resources (entrepreneurial orientation; risk taking and proactiveness), and their subsequent impact on company performance through the mediating role of innovativeness have not been thoroughly explored. This study aims to investigate in depth the effects of family branding and entrepreneurial orientation (risk taking and proactiveness), considered as resources with their respective components, on firm performance. The conceptual framework of investigated constructs is captured in Fig. 1.

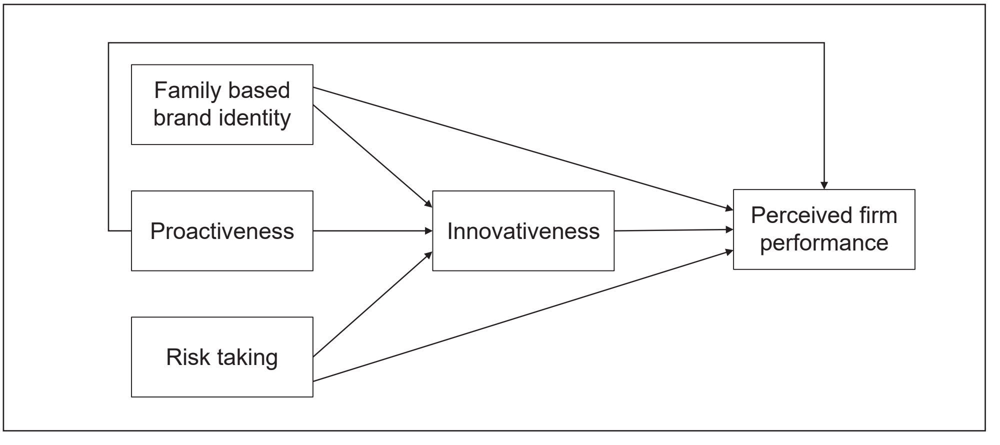

Fig. 1: Theoretical framework

Source: own

2 Research methodology

In this research, data were collected from human participants, and ethical approval was not required by our institution because the study posed minimal risk to participants. We informed all participants in advance that their information would be kept confidential and anonymous, and that they had the right to withdraw at any time during the survey.

Informed consent was obtained from all participants via a written electronic form prior to their involvement in the study, and we ensured that no deceptive practices were used. No participants were minors.

2.1 Questions topic

Family-based brand identity was measured using a four-item scale adapted from Craig et al. (2008). A pilot test with 30 managers from small family firms confirmed the clarity of the items. Respondents rated statements (e.g., "*Family ownership is highlighted in customer communications*") on a seven-point Likert scale (1 = strongly disagree to 7 = strongly agree): (i) "*I/We promote the fact that we are a family firm to our suppliers*"; (ii) "*I/We promote the fact that we are a family firm to our customers*"; (iii) "*I/We promote the fact that we are a family firm to our financiers*"; (iv) "*I/We include something about the fact that we are a family firm on our advertising materials, for example, on our letterhead, website, vehicles, etc.*"

The measure was adapted from the original scale developed by Miller (1983), which includes three items related to proactiveness. A Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) was used: (i) "*In dealing with competitors, our company typically initiates actions to which competitors then respond*"; (ii) "*In dealing with competitors, our company is very often the first firm to introduce new products/services, administrative techniques, operating technologies, etc.*"; (iii) "*In dealing with competitors, our company typically adopts a highly competitive, undo the competitors posture*."

The original risk-taking measure was developed by Covin and Slevin (1988). Responses were collected using a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree): (i) "*In general, our company has a strong proclivity (tendency) for high-risk projects with chances of very high returns*"; (ii) "*In general, our company believes that bold, wide-ranging actions are necessary to achieve the firm's objectives*"; (iii) "*When confronted with decision-making situations involving uncertainty, our company typically adopts a bold, aggressive posture to maximize the probability of exploiting potential opportunities.*"

This construct was originally developed by Miller (1983) and has been used in other studies (Kellermanns et al., 2012). The original scale included five items, but later, factor analysis led to the use of only two items to measure innovativeness (Kellermanns et al., 2012). In this study, all five originally developed items for measuring innovativeness are used to assess their mediating effects between the firm's resources and its performance in SFFs. Responses were collected on a scale ranging from 1 (strongly disagree) to 7 (strongly agree): (i) "*In our company, technical innovation based on research results, is readily accepted*"; (ii) "*In our company, management actively seeks innovative ideas*"; (iii) "*In our company, innovation is readily accepted in program/project management*"; (iv) "*In our company, people are rewarded for new ideas that work well*"; (v) "*In our company, innovation is perceived as constructive and is actively accepted*."

In this study, market share, performance rate, and profitability over the past three years were assessed using qualitative sentences on a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) to evaluate firm performance: (i) "*In comparison with our major competitors over the past three years, our company had a greater market share*"; (ii) "*In comparison with our major competitors over the past 3 years, our company has more performance rate*"; (iii) "*In comparison with our major competitors over the past 3 years, our company had higher profitability*."

3 Results

3.1 Sample and data collection

The present study concentrates on family-owned firms in Pakistan, employing an extensive dataset comprising 900 companies, verified up to the fourth quarter of 2022. A rigorous methodology was adopted to ensure data accuracy, involving the selection of samples from various sources. Notably, these samples were drawn from two distinct categories of authoritative sources: a record of SFFs associated with local trade associations and another record of family firms registered with the esteemed securities and exchange commission of Pakistan. It is important to highlight that these registries primarily contained essential contact information. Despite the utilization of these two initial sets of records as criteria, the sample size was constrained to 900 firms due to limitations in resources.

Data collection was strategically concentrated in Lahore and Faisalabad, the two principal industrial metropolises of Punjab province (Fig. 2). Located approximately 120 km apart, these cities form a dense industrial corridor that represents the core of the region's economic activity. Lahore, the provincial capital, and Faisalabad, collectively generate a significant portion of the national GDP (Government of Punjab, 2022). This selection captures the highest density of family firms in the country (SMEDA, 2022), thereby serving as a representative proxy for the broader SME sector in Pakistan. Efforts were made to ensure a balanced distribution of samples across these selected cities based on the registries mentioned earlier. The definition of family firms, as per (Sindhu et al., 2021), was employed, requiring at least one person from the family to be on the corporate board (involved in decision making), and the firm to have over 50% of shares owned by family members. Family members holding managerial positions were contacted, and 900 self-administered questionnaires accompanied by a cover letter were randomly selected. This was done using a non-probability convenience sampling method. Out of these, 600 questionnaires were physically distributed, followed by face-to-face interactions. The remaining 300 structured questionnaires were sent electronically via email or social media, using a Google Form, with reminder calls to encourage participation.

The convenience sampling technique was employed due to an unknown population size. Despite Urdu being Pakistan's national language and the presence of several regional languages, the survey was conducted in English. Although Urdu is the national language, English is the primary medium of instruction in higher education and the official language of business in Pakistan. Consequently, the use of English ensured that the instrument was administered to owners and top managers who possess the requisite professional proficiency, thereby maintaining the reliability of the data collection. Hair et al. (2018) advised eliminating responses with missing values to ensure data quality. Following this recommendation, 469 questionnaires with complete data were obtained, resulting in a 52.11% response rate. After excluding 9 incomplete or missing responses, a final sample of 460 firms (51.11%) was

Fig. 2: **Map of the study area (the industrial corridor of Punjab, Pakistan)**

Source: own based on https://www.google.com/maps selected for data analysis. The high response rate was attributed to the physical administration of questionnaires through direct interactions and a follow-up procedure involving reminder phone calls for online respondents. The demographic features of all respondents are presented in appendix (Tab. A1). Although detailed data on non-respondents were unavailable, the balanced distribution of respondents across firm size and age categories (see appendix – Tab. A1) suggests that the sample adequately represents the targeted population of SMEs in the region.

Measurement model

A confirmatory factor analysis was conducted on the final dataset to assess the discriminant validity of the five constructs: family branding, proactiveness, risk taking, innovativeness, and firm performance.

The preliminary measurement model consisted of 18 items assessing the five constructs mentioned, demonstrating adequate fit as all items exhibited loadings exceeding 0.7 onto their respective constructs (χ^2 (86) = 270.473, χ^2/df = 2.481, GFI = 0.927, AGFI = 0.898, CFI = 0.965, RMESA = 0.057, and SRMR = 0.039). In addition to that, two items of innovativeness and one item from family branding were deleted due to cross-loadings with other constructs to improve the model fit recommendations. This iterative elimination process culminated in a refined measurement model that adhered to both specific and overall goodness-of-fit benchmarks. (χ^2 (86) = 121.839, χ^2/df = 1.523, GFI = 0.966, AGFI = 0.949, CFI = 0.989, RMESA = 0.034, and SRMR = 0.0315) (see appendix – Tab. A2).

Further data validation, i.e., composite reliability (CR), average variance extracted (AVE), maximum shared variance (MSV), and inter construct correlations for each construct presented in appendix (Tab. A3), also achieves the threshold values (CR > 0.7, AVE > 0.5, MSV < AVE, √AVE > inter construct correlations), ensuring construct reliability, convergent and discriminant validity (Hair et al., 2018). Appendix (Tab. A3) presents all features of data including descriptive, reliability, and validity statistics.

3.2 Hypotheses testing

The structural model results (Tab. 1) reveal that intangible resources are critical drivers of SFF performance, collectively explaining 61.4% of the variance (R^2 = 0.61). Specifically, family-based brand identity emerged as the most dominant predictor (β = 0.642, p < 0.001). This indicates that in the Pakistani market, the trust, reputation, and distinctiveness associated with the family name are far more valuable assets than behavioral traits alone. Proactiveness also plays a vital role (β = 0.193, p < 0.001), suggesting that firms actively scanning for market opportunities significantly outperform reactive ones. In contrast, while risk-taking showed a positive effect, its direct impact was marginal compared to the other constructs (β = 0.064, p < 0.049), implying that mere boldness without strategic action yields limited direct returns. These results provide substantial support for hypotheses *H1–H3*.

Tab. 1: Results of hypotheses testing (direct effects)

Hypotheses	Independent variable	Beta	Sig.	*R*-squared	Observation
H1	Risk taking	0.064	0.049	61.40%	Accepted
H2	Proactiveness	0.193	0.000		Accepted
H3	Family branding	0.642	0.000		Accepted

Note: Dependent variable – firm performance.

Source: own

The bootstrap analysis (Tab. 2) confirms that innovativeness acts as a crucial conduit between firm resources and performance.

First, the relationship between family branding and performance is partially mediated by innovativeness (indirect effect = 0.389; 95% CI (0.323, 0.452)). This suggests a dual advantage: a strong family brand directly attracts customers through reputation, while simultaneously fostering an internal culture that encourages employees to share ideas and innovate. Similarly, proactiveness affects performance both directly and indirectly (indirect effect = 0.497; 95% CI (0.414, 0.581)). SFFs that anticipate future demands do not just gain a first-mover advantage; they actively translate these insights into tangible innovative outputs.

Perhaps the most critical finding concerns risk-taking. The analysis reveals full mediation (indirect effect = 0.124; 95% CI (0.020, 0.222)), as the direct path became non-significant in the presence of the mediator. This implies that for small family firms, risk-taking behavior is not inherently profitable. It only converts into improved performance when that risk is channeled specifically into innovative activities. In other words, blind risk-taking is ineffective; only risk-taking that fuels innovation drives value. Thus, hypotheses *H4–H6* are supported.

The confirmatory factor analysis and structural equation modeling results not only provide robust statistical evidence for the model's fit and hypothesized relationships but also offer important insights for understanding the drivers of performance in SFFs.

Interpretation of key findings

The standardized path coefficients indicate that family branding has the strongest direct effect on firm performance (β = 0.642, $p < 0.001$),

Tab. 2: Results of hypotheses testing (mediating effects)

Hypothesis	Mediating paths	Direct effects (LLCI; ULCI)	Indirect effects (LLCI; ULCI)	Observations
H4	Risk taking → performance (via INV)	0.208 (−0.0450: 0.0865)	0.1244 (0.0204: 0.2225)	Full
H5	Proactiveness → performance (via INV)	0.1539 (0.0798: 0.2279)	0.4969 (0.4140: 0.5807)	Partial
H6	Family branding → performance (via INV)	0.3817 (0.3158: 0.4475)	0.3891 (0.3237: 0.0452)	Partial

Source: own

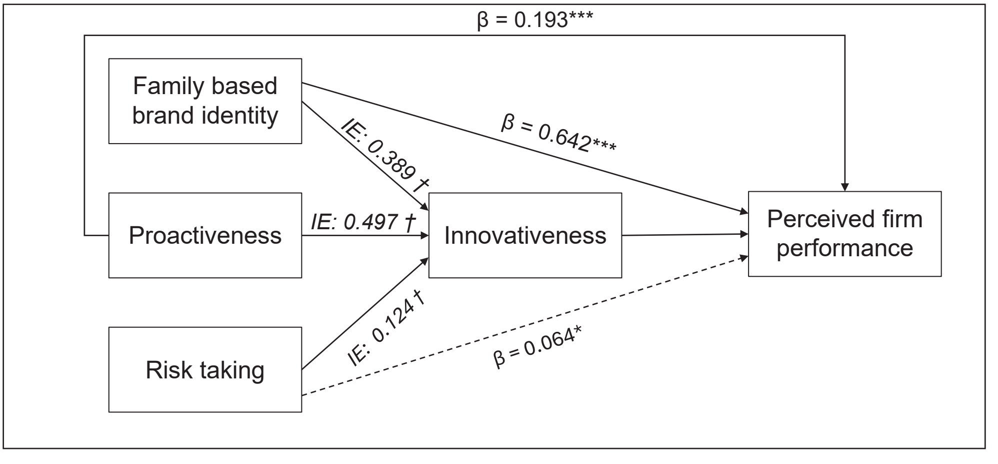

Fig. 3: Empirical results of the structural model

Note: β – standardized path coefficients for direct effects; IE – total indirect effect through Innovativeness; * $p < 0.05$; *** $p < 0.001$; † indicates significant indirect effect based on bootstrapping (LLCI/ULCI > 0).

Source: own

followed by proactiveness (β = 0.193, $p < 0.001$) and risk-taking (β = 0.064, $p < 0.049$). This ranking suggests that, in the context of Pakistani SFFs, building and leveraging a family-based brand identity represents the most powerful intangible resource for achieving competitive advantage and superior performance. Proactiveness contributes positively as well, emphasizing the need for firms to anticipate market changes and act on emerging opportunities. While risk-taking shows a significant effect, its smaller coefficient implies that boldness alone does not drive performance as strongly unless coupled with other strategies.

The model's *R*-squared value (61.4%) demonstrates that these intangible resources collectively explain a substantial proportion of the variance in perceived firm performance, underscoring their strategic importance in resource-constrained settings.

In practical terms, the size of the effect for family branding is considered large, while that for proactiveness is moderate and for risk-taking is small. These results highlight that, among the variables studied, family-based brand identity should be prioritized by SFF leaders aiming to strengthen their competitive position.

The empirical results provide strong support for the study's research objectives and hypotheses: confirming that family-based brand identity and entrepreneurial orientation (proactiveness and risk-taking) are valuable firm-specific resources in SFFs; demonstrating innovativeness as a key mediating mechanism; and empirically extending the RBV to the context of small family firms in emerging markets.

It is important to acknowledge that, as with all research, some limitations apply. The use of convenience sampling and the concentration on two major cities may limit the generalizability of the findings to all Pakistani SFFs or to family firms in other countries. Self-reported measures of firm performance, while common in management studies, are subject to potential respondent bias. These considerations suggest the value of future research employing more diverse sampling and objective performance data, as well as longitudinal designs to further validate the relationships observed.

4 Discussion

This study explores the relationships among family-based branding identity, entrepreneurial orientation, and the mediating role of innovativeness in SFFs. As previously discussed, family branding and entrepreneurial orientation are pivotal factors influencing a firm's innovativeness and overall performance (Jocic et al., 2023). Despite their significance in entrepreneurial firms, only a limited number of studies have considered these factors when examining determinants of innovativeness in SFFs.

Jocic et al. (2023) and Sherlock et al. (2023) contributed to examination of the impacts of family branding and entrepreneurial orientation as fundamental constructs, establishing a robust foundation for firm innovativeness and subsequent outcomes. However, these studies used global sample. The study aims address gap in specific region, specifically in Pakistan. The empirical findings confirm the central role of these constructs in shaping an innovative organization. Consequently, this research contributes to the theoretical understanding of the antecedents of family branding and entrepreneurial orientation, as well as the mediating influence of innovativeness, supported by concrete empirical evidence.

Furthermore, these results contextualize the resource-based view within an emerging market setting. Unlike studies on large, listed family firms in developed economies that often emphasize formal R&D and financial capital as primary performance drivers (Dabić et al., 2023), our findings highlight that SFFs in Pakistan rely heavily on strategic intangibles, which include not only brand reputation but also the effective implementation of knowledge management and adaptive managerial thinking (Bureš, 2006). In an environment characterized by institutional voids and limited formal financing, the trustworthiness and reputation associated with the family-based brand identity serve as a critical substitute for the structural resources possessed by larger conglomerates (Luo, 2019). This suggests that for SFFs, the family name is not merely a marketing tool, but a fundamental institutional asset that builds trust and legitimizes risk-taking and innovation (Beck & Prügl, 2018).

Additionally, this study in accordance with results of (Jocic et al., 2023; Sherlock et al., 2023) emphasizes the importance for non-listed SFFs to consistently prioritize and nurture entrepreneurial orientation in their organizational culture. This is crucial, as the inclination for proactiveness and risk taking may inherently favor SFFs over their larger counterparts. Top management should actively foster a culture of adventurous thinking and creativity by consistently promoting new ideas within the organization, thereby cultivating an environment conducive to innovativeness.

In the view of SFFs branding, this study can complement findings of Astner and Gaddefors (2025), who looked at SFFs branding from the perspective of the founder of SFF or add to Maldonado-Guzmán et al. (2023) where authors measured SFFs brand management. This research seeks to assess the applicability of the scale in the context of small-scale family firms. This implies that even in smaller firms, the extent to which a firm is dedicated to establishing its brand identity and entrepreneurial capacity can be empirically measured. In essence, this study underscores the need for ongoing attention to these factors, suggesting that maintaining a focus on brand identity and entrepreneurial capacity is relevant and measurable in the case of SFFs.

Implications. Through exploratory inquiries into proposed connections, the results indicate that SFFs should be inclined towards entrepreneurial activities and focus on establishing their family identity as a brand to seize more business opportunities which is align with Botero et al. (2018). Additionally, the findings support Fredyna et al. (2019) that the combination of entrepreneurial orientation and a strong sense of family branding can be bolstered in these firms through innovative behavior. These SFFs appear to possess the ability to generate intangible assets, particularly in the form of entrepreneurial orientation, through innovative culture and practices. The results further validate that both family branding and entrepreneurial orientation play pivotal roles as drivers of firm innovativeness. This, in turn, establishes innovativeness as a significant precursor to overall firm performance.

Many small entrepreneurial family firms operate with limited size and resources (Heider et al., 2022). Despite these constraints, they actively pursue success by leveraging core resources such as their brand identity, entrepreneurial orientation, and innovativeness. Consequently, the inclination to innovate, a crucial factor in maintaining competitiveness, may be more pronounced in such firms compared to their larger counterparts (Yin et al., 2023). It is recommended that top managers in SFFs devote considerable attention to enhancing innovativeness, with a specific focus on entrepreneurial practices, to achieve superior firm performance. The results indicate that family branding and entrepreneurial orientation positively influence a company's innovativeness, leading to improved overall performance. It appears that innovativeness plays a central role in integrating key factors that impact a firm's performance.

While companies may express a desire for innovation and breakthroughs in their research and development activities, progress may be limited if there is not a sufficient capacity for innovativeness (Paswan et al., 2021). This capacity, in turn, relies on the fundamental factors of entrepreneurial orientation and branding. Therefore, the findings suggest that both innovation managers and general managers should adopt a learning-oriented approach, with particular emphasis on developing both family branding and entrepreneurial orientation. This recommendation is particularly crucial as managers in many entrepreneurial firms are often the founders or part of the founding team of the company.

Future directions. Future studies should broaden the scope by examining the remaining dimensions of Entrepreneurial Orientation, specifically autonomy and competitive aggressiveness, which were not covered in this study. Furthermore, examining other potential moderators or mediators, such as family governance mechanisms or market turbulence, could provide a more nuanced understanding of the branding-performance-innovativeness link. It is also advisable for future research to comprehensively examine innovativeness in existing literature, giving particular attention to construct validity. Further, investigating other dimensions of risk taking, including uncertainties associated with venturing into uncharted territories, substantial investments in uncertain projects, and personal risks tied to unforeseeable professional challenges, can enrich future research in this domain (Ling-Zi et al., 2018; Tipu, 2017).

Conclusions

This study makes several novel contributions by positioning family-based brand identity and entrepreneurial orientation (proactiveness and risk taking) as critical intangible resources under the resource-based view, specifically in the context of small family firms operating in emerging markets. Through a rigorous SEM analysis of data from 460 privately held firms in Pakistan, it demonstrates that these family firm specific assets not only exert direct effects on perceived performance but also drive performance indirectly by fostering organizational innovativeness. By empirically validating innovativeness as a mediator, this research illuminates the mechanisms through which reputational and behavioral resources translate into competitive advantage, thereby enriching RBV theory and the family business literature. Crucially, the study clarifies the hierarchy of these resources in the Pakistani context: while family reputation serves as the dominant direct driver of success, risk-taking behavior is shown to be insufficient on its own, relying entirely on the mediating mechanism of innovativeness to generate value.

While prior research has established the general positive link between EO and performance (Rauch et al., 2009), this study moves beyond incremental validation by isolating the conditional nature of risk-taking. It demonstrates that in resource-constrained environments, risk-taking is not a direct driver of success but a latent resource that yields returns only when mediated by innovativeness. The theoretical implications of these findings include a deeper integration of family identity into resource-based frameworks and a more refined understanding of entrepreneurial orientation's role as a process capital in SFFs. From a managerial perspective, the results underscore the importance for family firm leaders to intentionally leverage their family name in stakeholder communications, institutionalize proactive decision making, and embrace calculated risks to cultivate an innovation-oriented culture.

From a practical perspective, the findings offer specific guidance for SFF owners and managers operating with limited budgets as summarized in Fig. 4:

(i) Leverage the family name as a strategic asset: owners should explicitly operationalize their family identity not just as a logo, but as a narrative tool. Marketing strategies should focus on storytelling that emphasizes family history, trustworthiness, and longevity to overcome the “liability of smallness” and institutional voids common in emerging markets;

(ii) Institutionalize innovation over blind risk-taking: since risk-taking alone does not drive performance, managers must avoid speculative ventures that lack an innovation component. Instead, SFFs should implement low-cost, structured innovation routines, such as regular “idea boxes” or family-employee brainstorming

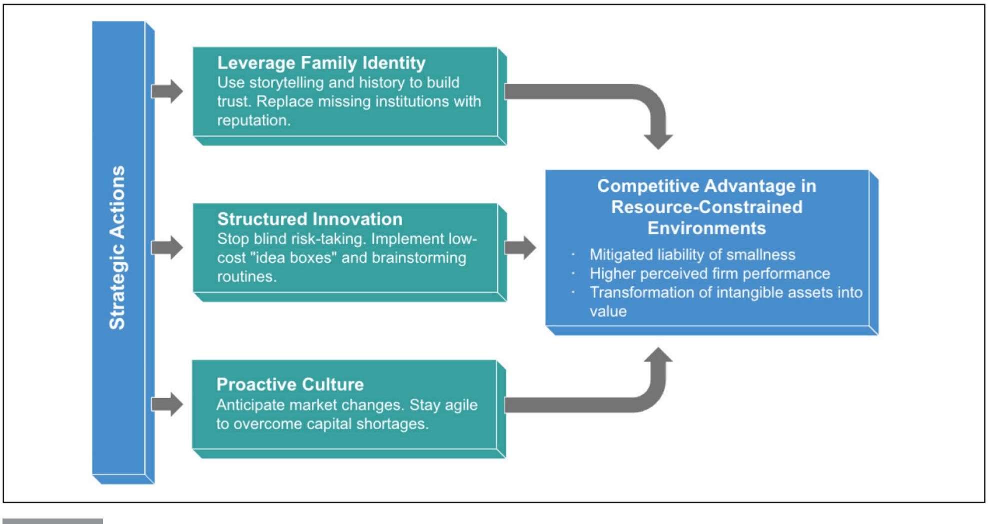

Fig. 4: **Strategic framework for managerial action in SFFs**

Source: own

sessions, to transform the firm's risk propensity into tangible process or product improvements. (iii) Foster a proactive culture: the study confirms that proactiveness impacts performance both directly and indirectly. Managers should encourage a culture where anticipating market changes is rewarded, ensuring the firm remains agile even without large capital reserves.

Despite its contributions, this research faces limitations. First, the cross-sectional design restricts causal inferences. Second, the survey was conducted in English. Although English is the standard language of business in Pakistan, this choice may have introduced a selection bias by underrepresenting micro-enterprises where owners possess lower English proficiency. Third, the reliance on self-reported perceptual measures of performance, while necessary due to the private nature of SFFs, may be subject to social desirability bias. Finally, while Lahore and Faisalabad represent the industrial core of the region, the concentration of sampling in these two metropolitan areas implies that the findings may not fully reflect the dynamics of family firms in rural areas or other emerging markets with different economic structures. Future studies should therefore employ longitudinal designs, strive to triangulate subjective responses with objective financial data, and incorporate additional dimensions of entrepreneurial orientation to further validate the model's external validity.

Acknowledgments: *The research has been partially supported by the Faculty of Informatics and Management UHK specific research project 2108 Addressing Current Challenges in the Field of Smart Cities and Cybersecurity.*

References

Arzubiaga, U., Maseda, A., & Iturralde, T. (2019). Entrepreneurial orientation in family firms: New drivers and the moderating role of the strategic involvement of the board. *Australian Journal of Management*, *44*(1), 128–152. https://doi.org/10.1177/0312896218780949

Astner, H., & Gaddefors, J. (2025). Founders and their brands: How founder identity matters in small firm branding. *Qualitative Market Research: An International Journal*, *28*(1), 101–121. https://doi.org/10.1108/QMR-06-2023-0085

Baquero, A. (2024). Linking green entrepreneurial orientation and ambidextrous green innovation to stimulate green performance: A moderated mediation approach. *Business Process Management Journal*, *30*(8), 71–98. https://doi.org/10.1108/bpmj-09-2023-0703

Barbera, F., Hasso, T., & Schwarz, T. V. (2023). Family governance and firm performance: Exploring the intermediate effects of family functioning and competitive advantage. *Journal of Family Business Management*, *13*(4), 1085–1103. https://doi.org/10.1108/jfbm-06-2022-0076

Beck, S., & Prügl, R. (2018). Family firm reputation and humanization: Consumers and the trust advantage of family firms under different conditions of brand familiarity. *Family Business Review*, *31*(4), 460–482. https://doi.org/10.1177/0894486518792692

Botero, I. C., Astrachan, C. B., & Calabrò, A. (2018). A receiver's approach to family business brands: Exploring individual associations with the term "family firm". *Journal of Family Business Management*, *8*(2), 94–112. https://doi.org/10.1108/jfbm-03-2017-0010

Buchanan, B., Martikainen, M., & Nikkinen, J. (2023). Family firm competitiveness and owner involvement. *Journal of Applied Accounting Research*, *24*(2), 260–281. https://doi.org/10.1108/jaar-11-2021-0317

Bureš, V. (2006). Knowledge management and its implementation. In J. A. M. Cordeiro, V. Pedrosa, B. Encarnacao, & J. Filipe (Eds), *WEBIST 2006: Proceedings of the Second International Conference on Web Information Systems and Technologies* (pp. 115–118). SciTePress.

Bureš, V., & Čech, P. (2007). Complexity of ambient intelligence in managerial work. In J. Hughes, R. Peiris, & P. Tyman (Eds), *Proceedings of the 12th Annual SIGCSE Conference on Innovation and Technology in Computer Science Education* (pp. 325–325). ACM. https://doi.org/10.1145/1268784.1268892

Buttar, H. M., & Kocak, A. (2011). The relationship between entrepreneurial orientation dynamic capabilities and firm performance: An exploratory study of small Turkish firms. *International Journal of Business and Globalisation*, *7*(3), 351. https://doi.org/10.1504/ijbg.2011.042063

Chatzoglou, P., & Chatzoudes, D. (2018). The role of innovation in building competitive advantages: An empirical investigation. *European Journal of Innovation Management*, *21*(1), 44–69. https://doi.org/10.1108/ejim-02-2017-0015

Covin, J. G., & Slevin, D. P. (1988). The influence of organization structure on the utility of an entrepreneurial top management style. *Journal of Management Studies*, *25*(3), 217–234. https://doi.org/10.1111/j.1467-6486.1988.tb00033.x

Covin, J. G., & Wales, W. J. (2012). The measurement of entrepreneurial orientation. *Entrepreneurship Theory and Practice*, *36*(4), 677–702. https://doi.org/10.1111/j.1540-6520.2010.00432.x

Craig, J. B., Dibrell, C., & Davis, P. S. (2008). Leveraging family-based brand identity to enhance firm competitiveness and performance in family businesses. *Journal of Small Business Management*, *46*(3), 351–371. https://doi.org/10.1111/j.1540-627x.2008.00248.x

Dabić, M., Vlačić, B., & Kovač, I. (2023). Guest editorial: The future of family business: Marketing challenges in times of crisis. *Journal of Family Business Management*, *13*(1), 1–6. https://doi.org/10.1108/jfbm-03-2023-159

Fang, H. “Chevy”, Kotlar, J., & Chrisman, J. J. (2025). Family control and sustainable competitive advantage. *Family Business Review*, *38*(2), 84–120. https://doi.org/10.1177/08944865251321881

Fredyna, T., Ruiz Palomo, D., & Dieguez Soto, J. (2019). Entrepreneurial orientation and product innovation. The moderating role of family involvement in management. *European Journal Of Family Business*, *9*(2), 128–145. https://doi.org/10.24310/ejfbejfb.v9i2.5392

Guo, Z., & Jiang, W. (2020). Risk-taking for entrepreneurial new entry: Risk-taking dimensions and contingencies. *International Entrepreneurship and Management Journal*, *16*(2), 739–781. https://doi.org/10.1007/s11365-019-00567-8

Hair, J. F., Harrison, D. E., & Risher, J. J. (2018). Marketing research in the 21st century: Opportunities and challenges. *Revista Brasileira de Marketing*, *17*(5), 666–699. https://doi.org/10.5585/bjm.v17i5.4173

Heider, A., Hülsbeck, M., & von Schlenk-Barnsdorf, L. (2022). The role of family firm specific resources in innovation: An integrative literature review and framework. *Management Review Quarterly*, *72*(2), 483–530. https://doi.org/10.1007/s11301-021-00256-3

Hügel, S. (2019). Reconsidering firm innovativeness. In *Innovation in service industries* (pp. 17–36). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-27179-4_3

Hult, G. T. M., Hurley, R. F., & Knight, G. A. (2004). Innovativeness: Its antecedents and impact on business performance. *Industrial Marketing Management*, *33*(5), 429–438. https://doi.org/10.1016/j.indmarman.2003.08.015

Hussain, S. A. (2017). How to make knowledge resources valuable: An insight of knowledge integration process analysis. *VINE Journal of Information and Knowledge Management Systems*, *47*(1), 42–54. https://doi.org/10.1108/vjikms-01-2016-0002

Huybrechts, J., Voordeckers, W., Lybaert, N., & Vandemaele, S. (2011). The distinctiveness of family-firm intangibles: A review and suggestions for future research. *Journal of Management & Organization*, *17*(2), 268–287. https://doi.org/10.5172/jmo.2011.17.2.268

Jocic, M. R., Morris, M. H., & Kuratko, D. F. (2023). Familiness and innovation outcomes in family firms: The mediating role of entrepreneurial orientation. *Journal of Small Business Management*, *61*(4), 1345–1377. https://doi.org/10.1080/00472778.2020.1861284

Kellermanns, F. W., Eddleston, K. A., & Zellweger, T. M. (2012). Article commentary: Extending the socioemotional wealth perspective: A look at the dark side. *Entrepreneurship Theory and Practice*, *36*(6), 1175–1182. https://doi.org/10.1111/j.1540-6520.2012.00544.x

Kreiser, P. M., & Davis, J. (2010). Entrepreneurial orientation and firm performance: The unique impact of innovativeness, proactiveness, and risk-taking. *Journal of Small Business & Entrepreneurship*, *23*(1), 39–51. https://doi.org/10.1080/08276331.2010.10593472

Kyrgidou, L. P., & Spyropoulou, S. (2013). Drivers and performance outcomes of innovativeness: An empirical study. *British Journal of Management*, *24*(3), 281–298. https://doi.org/10.1111/j.1467-8551.2011.00803.x

Lee, T., & Chu, W. (2013). How entrepreneurial orientation, environmental dynamism, and resource rareness influence firm performance. *Journal of Management & Organization*, *19*(2), 167–187. https://doi.org/10.1017/jmo.2013.33

Ling-Zi, Y., Shu, L., & Zhu-Yuan, L. (2018). New avenues for the development of domain-specific nature of risky decision making. *Advances in Psychological Science*, *26*(5), 928. https://doi.org/10.3724/sp.j.1042.2018.00928

Luo, X. R. (2019). Family business and SMEs. *Management and Organization Review*, *15*(1), 211–212. https://doi.org/10.1017/mor.2019.7

Madanoglu, M., Memili, E., & De Massis, A. (2020). Home-based family firms, spousal ownership and business exit: A transaction cost perspective. *Small Business Economics*, *54*(4), 991–1006. https://doi.org/10.1007/s11187-018-00131-8

Maldonado-Guzmán, G., Molina-Morejón, V. M., & Juárez-del Toro, R. (2023). Brand management measurement in Mexican small firms: Empirical evidence. *Advances in Management and Applied Economics*, 123–140. https://doi.org/10.47260/amae/1357

Melewar, T. C., Gambetti, R. C., & Martin, K. D. (2014). Special issue on: Managing intangible ethical assets: Enhancing corporate identity, corporate brand, and corporate reputation to fulfill the social contract. *Business Ethics Quarterly*, *24*(3), 504–506. https://doi.org/10.1017/s1052150x00006291

Miller, D. (1983). The correlates of entrepreneurship in three types of firms. *Management Science*, *29*(7), 770–791. https://doi.org/10.1287/mnsc.29.7.770

Mishra, C. S. (2021). Does managerial ability drive firm innovativeness? *IEEE Transactions on Engineering Management*, *68*(4), 1139–1154. https://doi.org/10.1109/tem.2019.2923892

Morris, M. H., Kuratko, D. F., & Covin, J. G. (2010). *Corporate entrepreneurship & innovation*. Cengage Learning.

Paswan, A. K., Guzmán, F., & Pei, Z. (2021). Innovation-branding: Should all firms be equally ambidextrous? *Journal of Product & Brand Management*, *30*(5), 754–767. https://doi.org/10.1108/JPBM-07-2019-2476

Rauch, A., Wiklund, J., Lumpkin, G. T., & Frese, M. (2009). Entrepreneurial orientation and business performance: An assessment of past research and suggestions for the future. *Entrepreneurship Theory and Practice*, *33*(3), 761–787. https://doi.org/10.1111/j.1540-6520.2009.00308.x

Riivari, E., & Lämsä, A.-M. (2014). Does it pay to be ethical? Examining the relationship between organisations' ethical culture and innovativeness. *Journal of Business Ethics*, *124*(1), 1–17. https://doi.org/10.1007/s10551-013-1859-z

Rua, O., França, A., & Fernández Ortiz, R. (2018). Key drivers of SMEs export performance: The mediating effect of competitive advantage. *Journal of Knowledge Management*, *22*(2), 257–279. https://doi.org/10.1108/jkm-07-2017-0267

Sarsiti, S., & Minarni, E. (2024). Entrepreneurial orientation (EO) and performance among micro, small and medium enterprises (MSMEs) (from an RBV perspective). *DiE: Jurnal Ilmu Ekonomi dan Manajemen*, *15*(2), 227–241.

Sherlock, C., Dibrell, C., & Memili, E. (2023). The impact of family commitment on firm innovativeness: The mediating role of resource stocks. *Journal of Family Business Strategy*, *14*(3), 100523. https://doi.org/10.1016/j.jfbs.2022.100523

Sindhu, M., Saleem, I., & Arshad, M. (2021). When do family brand personalities lead to brand loyalty? A study of family-owned fashion retailers in Pakistan. *Global Business and Organizational Excellence*, *40*(6), 6–16. https://doi.org/10.1002/joe.22133

SMEDA. (2022). *SMEDA annual report (brief): Fiscal year of 2021–22*. https://smeda.org/phocadownload/Publicatoins/SMEDA%20Annual%20Report%202021-22%20Brief%20-%20Infograph.pdf

Smith, C., Rondi, E., De Massis, A., & Nordqvist, M. (2024). Rising every time we fall: Organizational fortitude and response to adversities. *Journal of Management*, *50*(5), 1865–1910. https://doi.org/10.1177/01492063231164969

Sunday Aneke, F., & Ja'afaru Garba, S. (2024). Unlocking competitive advantage: Exploring the role of entrepreneurial orientation on small and medium enterprises in Kogi State. *SPECTRUM Journal of Social Science*, *1*(3), 79–94. https://doi.org/10.61552/sjss.2024.03.003

Tipu, S. A. A. (2017). Entrepreneurial risk taking: Themes from the literature and pointers for future research. *International Journal of Organizational Analysis*, *25*(3), 432–455. https://doi.org/10.1108/ijoa-08-2015-0898

Valentin, G., & Călin, D. (2024). The role of organizational culture in driving innovation: A study of contemporary business practice. *Acta Universitatis Cibiniensis. Technical Series*, *76*(1), 46–54. https://doi.org/10.2478/aucts-2024-0007

Wiid, J., Cant, M., & Senooane, B. (2025). An inside-out model of retail SME brand orientation for an emerging market. *Management Dynamics in the Knowledge Economy*, *13*(1), 1–19. https://doi.org/10.2478/mdke-2025-0001

Yin, Y., Crowley, F., Doran, J., Du, J., & O'Connor, M. (2023). Research and innovation and the role of competition in family owned and managed firms. *International Journal of Entrepreneurial Behavior & Research*, *29*(1), 166–194. https://doi.org/10.1108/IJEBR-12-2021-1031

Appendix

Tab. A1: Demographic information

	Frequency	Percentage (%)
Nature of the business		
Manufacturing	354	77
Services	106	23
Age of the firm		
1–9 years	315	68.5
10–19 years	94	20.4
20–50 years	37	8
50–75 years	14	3
Work experience of respondents		
1–10 years	288	62.6
11–20 years	61	13.3
21–30 years	66	14.3
More than 30 years	45	9.8
Number of employees		
Less than 10	75	16.3
10–30 employees	198	43
30–100 employees	132	28.7
100–500 employees	45	9.8
Above 500 employees	10	2.2
Position held by the respondents		
Top-level management (e.g., CEO, managing director, chairman)	296	64.3
Middle Level (e.g., department head, senior managers)	128	27.8
Lower-level management (e.g., line managers)	36	7.8

Source: own

Tab. A2: Measurement model

Measurement items	Items	Loadings	CR	AVE
Risk taking (RTT)	RT1	0.877	0.886	0.721
	RT2	0.809		
	RT3	0.859		
Proactiveness (PRR)	PR1	0.819	0.803	0.803
	PR2	0.958		
	PR3	0.906		
Family branding (FBB)	FB2	0.932	0.959	0.886
	FB3	0.956		
	FB4	0.935		
Innovation (INVV)	INV2	0.831	0.909	0.909
	INV3	0.907		
	INV5	0.892		
Perceived firm performance (PFPP)	PFP1	0.934	0.965	0.965
	PFP2	0.961		
	PFP3	0.955		

Source: own

Tab. A3: Discriminant and convergent validity

	CR	AVE	MSV	MaxR(H)	INVV	FBB	PRR	RTT	PFPP
Innovativeness	0.909	0.770	0.640	0.915	0.877				
Family branding	0.959	0.886	0.582	0.960	0.665***	0.941			
Proactiveness	0.924	0.803	0.354	0.947	0.585***	0.595***	0.896		
Risk taking	0.886	0.721	0.013	0.890	0.116*	0.068	0.042	0.849	
Perceived firm's performance	0.965	0.902	0.640	0.967	0.800***	0.763***	0.586***	0.115*	0.950

Note: CR – composite reliability; AVE – average variance extracted; MSV – maximum shared variance; MaxR(H) – maximal reliability; INVV – innovativeness; FBB – family branding; proactiveness (PRR); RTT – risk taking; PFPP – perceived firm's performance.

Source: own

Explaining cross-country differences in firms' innovation outcomes: The role of corruption, labor regulation, productivity and gender in the EU-27

Nikola Soukupova[1]*, Anna Nemcova*[2]

[1] *University of Pardubice, Faculty of Economics and Administration, Institute of Economy and Management, Czech Republic, ORCID: 0000-0002-2462-5641, nikola.soukupova@upce.cz (corresponding author);*

[2] *University of Pardubice, Faculty of Economics and Administration, Institute of Economy and Management, Czech Republic, ORCID: 0000-0002-8316-5737, anna.nemcova@upce.cz.*

Abstract: *Innovation outcomes are shaped by both institutional conditions and firms' internal capabilities, yet prior research often examines these dimensions separately and rarely distinguishes between different types of innovation. This limits understanding of how external barriers and organizational resources jointly influence incremental versus market-novel innovation. To address this gap, the aim of this paper is to examine how institutional constraints, specifically perceived corruption and perceived labor regulation, and firm-level capabilities, measured by labor productivity growth, gender diversity in top management, and firm size (as a control), jointly shape firms' innovation performance. Grounded in a unified institutional and resource-based view (RBV) framework, we test how de facto institutional frictions and internal capabilities differentially relate to incremental versus market-novel innovation. Using harmonized World Bank Enterprise Survey data from 2019–2023 covering EU-27 countries (5,534 firms), we estimate two ordinary least squares (OLS) models to explain cross-country differences in: (i) the share of firms introducing new products or services (incremental innovation); and (ii) the share of firms whose innovations are new to the main market (market-novel innovation). All estimates are at the country level (EU-27; N = 27) and are interpreted as between-country associations in the national share of firms innovating. Accordingly, we interpret results as cross-country patterns and avoid firm-level causal claims. The results reveal a clear divergence in determinants across innovation types. Perceived corruption significantly reduces incremental innovation but does not affect market novelty. Labor regulations do not influence incremental innovation yet show a positive association with market novelty, consistent with a compliance-induced innovation mechanism. Productivity growth displays a dual effect: it suppresses incremental innovation but enables market novelty through greater organizational slack. Gender diversity in top management and firm size further strengthen market-novel innovation, while showing no effect on incremental outcomes. The study contributes to institutional theory by demonstrating that governance weaknesses primarily constrain routine innovation and refines RBV by showing that internal capabilities – productivity, organizational scale and leadership diversity are decisive drivers of market-novel innovation. By linking institutional environments with firm-level resources, the study provides an integrated explanation of innovation heterogeneity across EU economies.*

Keywords: *Innovation, corruption, labor regulation, labor productivity, gender, European firms.*

JEL Classification: *O30, O31, O38, D22, L25.*

APA Style Citation: Soukupova, N., & Nemcova, A. (2026). Explaining cross-country differences in firms' innovation outcomes: The role of corruption, labor regulation, productivity and gender in the EU-27. *E&M Economics and Management*, *29*(2), 141–155. https://doi.org/10.15240/tul/001/2026-2-010

Introduction

Innovation is a key driver of productivity growth and long-term competitiveness (Hall, 2011). Firms' innovation decisions are shaped both by their internal capabilities and by the institutional environments in which they operate (Donges et al., 2023). Institutional theory posits that organizations operate within formal and informal rules that structure incentives, constraints and legitimacy pressures (North, 1990; Scott, 2014). These institutional arrangements influence access to resources and shape firms' perceptions of risk and expected innovation returns. Recent work highlights the interaction between institutional environments and firm-level resource endowments in shaping innovation outcomes (Natera & Castellacci, 2021; Peng et al., 2020). Complementing this perspective, the resource-based view (RBV) emphasizes firms' unique and hard-to-imitate resources and dynamic capabilities as foundations of competitive advantage (Barney, 1991; Schilke et al., 2018; Teece, 2014). Together, these two perspectives indicate that institutions define the external opportunity structure, while internal capabilities determine how effectively firms can exploit it. Consistent with this dual perspective, we first focus on corruption as a salient institutional friction because it directly raises transaction costs and reduces the appropriability of innovation returns, thereby reshaping firms' expected payoffs from innovation (Cuervo-Cazurra, 2016; North, 1990; Scott, 2014).

Against this backdrop, we begin by examining corruption as a primary institutional barrier. A broad empirical literature shows that corruption reduces the likelihood of introducing new products or services by increasing uncertainty and transaction costs (Cuervo-Cazurra, 2016; Ellis et al., 2020; Lee et al., 2020). Weak institutional environments further discourage both incremental and radical innovation (Fisman et al., 2024; OECD, 2023) and undermine innovation quality, reducing the number and impact of green patents (Quan et al., 2023). Political connections also steer firms toward safer, less transformative innovation trajectories (Zhong & Zheng, 2025).

The impact of labor market regulation on innovation remains theoretically ambiguous. Employment protection may encourage investment in firm-specific skills, yet rigid rules can hinder flexibility and slow resource reallocation (Acharya et al., 2013). Firms' perceptions of regulatory stringency and administrative burden often shape their responses as much as statutory provisions (The World Bank, 2019–2023). From an RBV perspective, firm-level capabilities, such as productivity, managerial composition and organizational scale, affect absorptive capacity and influence how firms translate external conditions into innovation (Stříteská et al., 2016; Teece, 2018).

Despite extensive research, important gaps remain. Institutional and firm-level determinants are rarely examined within a unified framework. Less is known about how perceived corruption and perceived labor regulation influence innovation outcomes, especially using harmonized, cross-country firm-level data. Moreover, many cross-country studies rely on input-based indicators (R&D, patents) rather than actual product and market-level innovation outcomes. The aim of this paper is to explain cross-country variation in the share of firms that: (i) introduce new products or services; and (ii) bring innovations new to the main market, focusing on de facto institutional constraints (perceived corruption, perceived labor regulation) and on firm capabilities measured as countrylevel aggregates (labor productivity growth, the share of firms with a woman in top management, and average firm size as a control).

Drawing on World Bank Enterprise Surveys (The World Bank, 2019–2023), the analysis integrates institutional and firm-level factors to provide a multi-level explanation of innovation performance. Our selection of variables follows the joint logic of institutional theory and the resource-based view. Perceived corruption and perceived labor regulation capture economy-wide rules that shape risk, appropriability and adjustment costs (Acharya et al., 2013; Cuervo-Cazurra, 2016; North, 1990; Scott, 2014). Labor productivity growth, gender diversity in top management and firm size proxy firms' resource endowments, absorptive capacity and coordination capabilities (Barney, 1991; Hall et al., 2009; Leiponen & Helfat, 2010; Teece, 2014). We acknowledge other relevant determinants such as access to finance, export intensity and competitive pressure; these are conceptually important but lie outside our focal model to maintain parsimony and a clear mapping between theory and measurement given the cross-country design. Firm size is retained as a control to net out scale-related heterogeneity, and any significant association is interpreted in this role rather than as a primary explanatory mechanism (Hall et al., 2009; Leiponen & Helfat, 2010). The remainder of the paper develops the hypotheses, presents the data and methods, reports and discusses the findings and concludes with theoretical and practical implications.

1 Theoretical background and hypothesis development

Perceived corruption constitutes a major institutional barrier for firms, reducing the likelihood of introducing new products or services (De Rosa et al., 2015; Lee et al., 2020). Corrupt environments increase the transaction costs of innovation. Instead of enhancing internal productivity, managers must devote time and resources to negotiations with officials, which undermines governance systems (Bó & Rossi, 2007). At the same time, the appropriability of future returns from innovation declines, as access to licenses and intellectual property protection is frequently conditioned on informal payments. Higher costs and lower expected returns, therefore, prompt firms to delay or abandon innovation projects (Bó & Rossi, 2007). From the perspective of institutional theory, corruption distorts the rules of the game, erodes trust in formal institutions, and systematically weakens firms' incentives to engage in uncertain innovation projects (North, 1990; Peng et al., 2020; Scott, 2014). By raising transaction costs and lowering appropriability, corruption undermines the institutional foundations of innovative activity (Bănică et al., 2024). This theoretical perspective implies that innovation depends not only on the availability of financial or technological resources but also on the predictability and fairness of the institutional environment.

Empirical studies on firms in Central and Eastern Europe confirm that higher regional corruption reduces the likelihood of both product and process innovations, even after controlling for regulatory and financial constraints (De Rosa et al., 2015; Lee et al., 2020). Comparable evidence beyond CEE also shows that corruption depresses not only the quantity but the quality of innovative output in firms (e.g., among U.S. firms), reinforcing the institutional mechanism above (Ellis et al., 2019). More recent research further indicates that corruption also affects innovation quality. Quan et al. (2023) found that political corruption systematically decreases the number of green patents and their citation impact, implying weaker technological and economic contributions. At the same time, work on political ties suggests that when anti-corruption shocks sever connections, firms may temporarily redirect effort toward patenting, but this does not necessarily translate into higher grant rates, consistent with heightened uncertainty and short-term signaling (Zhong & Zheng, 2025). At the level of European regions, Bănică et al. (2024) confirm that institutional quality, particularly control of corruption and rule of law, is crucial for sustainable innovation performance and the generation of high-quality knowledge.

Business surveys reveal similar conclusions. A higher frequency of "facilitation payments" is associated with a lower probability of technological innovation (Shabir et al., 2020). The negative consequences of corruption are most pronounced in industries characterized by frequent interactions with public administration and complex licensing procedures, where informal payments are more likely (Yakovlev & Zhuravskaya, 2013). In these sectors, owners often prioritize managers with political connections over technically competent ones, which reduces productivity and innovation capacity (Mironov, 2015). This aligns with the findings of Zhong and Zheng (2025), who show that political connections constrain both the scope and radicalness of innovation, especially in contexts with weak rule of law.

Weak contractual institutions combined with corruption further hinder technological progress, particularly in industries dependent on contract-based inputs. Bănică et al. (2024) add that in the context of green and digital transformation, institutional quality is essential not only for the volume but also for the strategic orientation of innovation. Importantly, firms' day to day experience of the business environment often reflects "deals rather than rules", meaning that perceived obstacles capture de facto frictions that formal indicators may miss (Hallward-Driemeier & Pritchett, 2015). From a policy perspective, this implies that financial incentives, such as R&D tax credits, may not suffice to stimulate innovation unless accompanied by systematic reduction of regulatory burdens, which, in fact, increased after EU accession (Alfano et al., 2020). Strengthening the rule of law and enhancing anti-corruption enforcement thus appear to be critical prerequisites for boosting productivity and innovation activity of European firms (Lee et al., 2020). Moreover, Dokas et al. (2023) demonstrate that innovation can mitigate the negative impact of corruption on economic growth, particularly in advanced economies.

Based on the above arguments, we hypothesize that:

H1a: Higher perceived corruption as a major obstacle to doing business reduces the likelihood of incremental innovation (introduction of new products/services) by increasing transaction costs and lowering appropriability of returns.

H1b: Higher perceived corruption reduces the likelihood of market-novel innovation, as systemic uncertainty discourages high-risk projects.

Labor regulations, particularly employment protection legislation (EPL), exert an ambivalent influence on firms' innovation performance. On the one hand, they prevent arbitrary dismissals, mitigating the hold-up problem and encouraging employees to invest in firm-specific knowledge – a key input for the emergence and development of innovations (Acharya et al., 2013). On the other hand, they raise restructuring costs and reduce flexibility in reallocating resources, potentially lowering innovation productivity, especially in industries requiring rapid adaptation and short learning cycles (Bai et al., 2020). The overall effect of EPL is therefore highly context-dependent – shaped by the technological intensity of industries, the degree of competitive pressure, and firms' strategic orientation (Acharya et al., 2013). Institutional perspectives suggest that regulatory frameworks shape the opportunity set of firms by defining the balance between flexibility and stability (Natera & Castellacci, 2021; Scott, 2014). When regulations are perceived as excessive, they impose institutional pressures that discourage experimentation and risk-taking; when perceived as stabilizing, they provide legitimacy and a foundation for long-term investments.

In recent years, greater attention has been devoted to the perceptual dimension of regulation, which captures how firms subjectively assess its stringency and impact. This approach encompasses not only the formal legislative framework but also its enforceability, administrative complexity, and the quality of supervisory mechanisms (The World Bank, 2019–2023). When firms perceive labor regulations primarily as a bureaucratic or financial burden, they may pursue radical product innovations as a new source of competitive advantage (Bassanini et al., 2009; Gao & Anwar, 2024). Conversely, firms that perceive regulations as a stabilizing framework for employee relations may focus on incremental improvements of their existing offerings (Filippetti & Guy, 2020). The negative effects of rigid labor regulations on innovation can be alleviated by factors such as high-quality human resource management, strong innovation culture, or the availability of flexible employment arrangements, which enable firms to benefit from workforce stability without sacrificing adaptability (Bassanini et al., 2009; Gao & Anwar, 2024).

Based on the above arguments, we hypothesize that:

H2a: When perceived as a significant barrier, labor regulations reduce incremental innovation by limiting firms' short-term flexibility.

H2b: When perceived as a significant barrier, labor regulations reduce market-novel innovation by constraining strategic reallocation capacity.

The 2023 Survey of Adult Skills (PIAAC) confirmed that cross-country differences in workers' competencies significantly contribute to productivity gaps. At the industry level, a higher average skill level among employees is systematically associated with higher productivity, potentially explaining up to one-quarter of cross-country differences in industrial productivity (Hanushek & Woessmann, 2015; OECD, 2023). A greater share of highly skilled employees also facilitates innovation diffusion and accelerates technology adoption, thereby strengthening firms' ability to respond to competitive pressures (Bartelsman et al., 2013).

Productivity is also indirectly influenced by the effective allocation of human capital. Skills mismatch, when workers' skills do not correspond to job requirements, leads to inefficient resource utilization and lowers aggregate productivity, particularly in R&D-intensive industries (Quintini, 2011). Empirical estimates suggest that approximately 10%–12% of productivity differences across countries can be attributed to the distribution of skilled workers across firms (OECD, 2023). While overqualification may bring short-term benefits at the firm level, at the macroeconomic level such misallocation impedes the expansion of more productive firms and slows aggregate productivity growth (Moretti, 2021).

Historically, Europe has lagged behind the United States in effectively reallocating human capital to the most productive firms, which has resulted in a slower pace of technological renewal. Key factors include lower entry of new dynamic firms, weaker growth of large leaders, and the persistence of less productive firms in the market (Bartelsman et al., 2013). Between 2003 and 2023, the average annualized growth of total factor productivity (TFP) among European firms in non-technological industries reached only 0.9%, compared to 2.6% in the United States; in the technology sector, European productivity declined by 0.3% per year, while U.S. productivity grew by 1.5% (OECD, 2023). Similar gaps are evident in R&D investment: U.S. technology firms allocate on average 12% of revenues to R&D, while European firms only 3%–4%, contributing to the widening gap in both the quantity and quality of patents (OECD, 2023). From the RBV, labor productivity represents a fundamental internal capability that enables firms to reallocate resources, absorb external knowledge, and pursue higher-risk projects (Barney, 1991; Schilke et al., 2018; Teece, 2018). Productivity growth not only reflects efficiency improvements but also provides the internal slack required for long-term investments in uncertain and market-novel innovations. Firms with higher productivity are therefore better positioned to exploit dynamic capabilities and transform resource advantages into superior innovation outcomes (Natera & Castellacci, 2021). Consistent with this capability view, micro-evidence from European SMEs links stronger innovation performance to firm-level productivity and associated absorptive capacity (Hall et al., 2009).

Building on the RBV mechanism whereby productivity growth creates slack and strengthens absorptive capacity, we hypothesize that:

H3a: Growth in labor productivity negatively correlates with incremental innovation, as efficiency improvements substitute for minor product changes.

H3b: Growth in labor productivity positively correlates with market-novel innovation, as enhanced efficiency creates resource slack for ambitious projects.

Research indicates that gender diversity in top management teams (TMT) enhances firms' innovation capabilities and strategic adaptability. Diverse perspectives and experiences improve the ability to identify new market opportunities and make informed decisions, particularly in environments of rapid technological change (Jeong & Harrison, 2017; Lee & Chung, 2022). Female leadership styles, characterized by participative decision-making and open communication, foster knowledge sharing and collaboration across organizational units (Van Knippenberg et al., 2004). Such a culture facilitates the transformation of new ideas into innovative products and services (Tang et al., 2021) and can reduce risk aversion in long-term project investments (Lee & Chung, 2022).

Empirical studies in European contexts confirm that firms with higher representation of women in top management are not only more likely to introduce innovations new to the main market, but also more effective in leveraging knowledge from external networks and research institutions, thereby strengthening their long-term innovation performance (Capozza & Divella, 2024; Soukupova et al., 2024). From the RBV, gender diversity constitutes a valuable intangible resource that enhances the cognitive and relational capital of the firm (Barney, 1991; Teece, 2014). Aligned with this mechanism, evidence shows that the performance effect of women in top management strengthens when a firm's strategy is innovation-oriented, and that gender-diverse R&D teams are more likely to produce radical, higher-novelty outcomes (Capozza & Divella, 2024; Dezső & Ross, 2012).

Based on the above arguments, we hypothesize that:

H4a: The presence of a woman in top management does not significantly affect incremental innovation.

H4b: The presence of a woman in top management increases the likelihood of market-novel innovation by enhancing decision-making diversity and openness to external knowledge.

The size of the workforce, measured by the number of full-time equivalent (FTE) employees, is a significant determinant of firms' innovation capacity. Larger firms typically possess more extensive resources, specialized teams, and more advanced R&D infrastructure, which increases their ability to implement complex innovation projects and commercialize outcomes (Leiponen & Helfat, 2010). From a resource-based view, firm size represents a structural capability that enhances the breadth of available knowledge, complementary skills, and absorptive capacity (Beck et al., 2022).

However, large organizational structures may also lead to bureaucratic rigidities, slower decision-making, and reduced adaptability in rapidly changing technological environments. Empirical studies confirm that the optimal effect of firm size on innovation depends on industry structure, technological intensity, and strategic orientation (Boly et al., 2022). Evidence from European contexts suggests that in high-tech sectors, the positive relationship between workforce size and innovation performance tends to be stronger, whereas in traditional industries, smaller and more agile firms often play a more significant role (Hall et al., 2009). This relationship may also be moderated by the quality of human capital – firms with a higher share of highly skilled employees achieve greater returns from R&D investment regardless of total FTE numbers (Boly et al., 2022). In this study, firm size is included as a control variable to account for differences in firms' resource endowment and organizational capacity that may influence innovation performance. From an RBV perspective, size captures a firm's potential to mobilize resources and coordinate knowledge, while also reflecting organizational constraints that shape its adaptive capacity. In line with this role, we treat firm size as a control rather than a focal explanatory variable. When its coefficient is statistically significant, we interpret it as a scale effect that nets out structural heterogeneity, consistent with prior evidence on size-innovation relationships (Hall et al., 2009; Leiponen & Helfat, 2010).

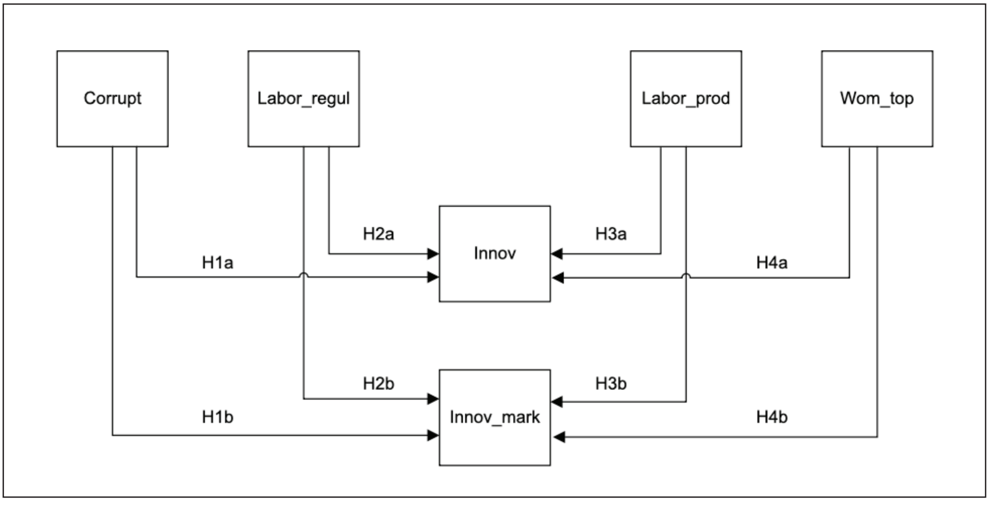

Fig. 1: Conceptual framework

Source: own

Based on the reviewed literature and formulated hypotheses, the conceptual framework (Fig. 1) integrates institutional and organizational determinants of innovation. Building on institutional theory (Natera & Castellacci, 2021; Peng et al., 2020), the model positions perceived corruption and labor market regulations as external constraints shaping firms' incentives and capacity to innovate. From a RBV (Barney, 1991; Schilke et al., 2018; Teece, 2014), labor productivity and gender diversity in top management are conceptualized as internal capabilities that enhance firms' ability to generate and commercialize new ideas. Innovation is operationalized in two complementary dimensions: general innovation outcomes (*Innov*) and market-oriented innovation (*Innov_mark*). Eight hypotheses capture these relationships, testing the effects of corruption, labor market regulation, productivity, and women's representation in top management. The framework underscores that innovation outcomes are not driven by institutional or organizational factors alone but by their interaction, offering a systematic basis for understanding how governance quality, labor market efficiency, and inclusive leadership jointly shape firms' competitiveness.

2 Research methodology

2.1 Data source

The empirical analysis is based on the World Bank Enterprise Surveys (WBES) (The World Bank, 2019–2023) conducted between 2019 and 2023. The WBES provides nationally representative evidence on firms in the private manufacturing and services sectors, using a standardized methodology that ensures cross-country comparability. For this study, data from 27 EU member states are used, covering 5,534 surveyed firms in total. Each country is surveyed in a specific year within this period, meaning that the dataset represents a cross-sectional snapshot rather than a balanced panel. Unit of analysis and aggregation. The unit of analysis is the country. Firm-level responses from the WBES are transformed into weighted countrylevel indicators using the survey weights provided by the World Bank. For binary/categorical items (e.g., introduced an innovation in the last three years; innovation new to the main market; corruption as the biggest obstacle; labor regulation as a major/very severe obstacle; woman top manager; provision of formal training), we compute weighted proportions. For continuous attributes (e.g., average firm size), we compute weighted means; national labor-productivity growth is aligned to the survey reference window. This design yields one observation per country (EU-27; $N = 27$). Consequently, estimates reflect between-country associations in the national share of firms innovating; we do not draw firm-level causal inferences. Given the small-N setting, we privilege parsimonious specifications and heteroskedasticity-robust standard errors. The WBES offers harmonized and representative measures that make it well suited for cross-country research in a European setting. The shared EU policy framework focused on competitiveness, digital transformation, and sustainable growth further enhances comparability while still allowing for meaningful differences in national business environments.

2.2 Description of variables

Two dependent variables capture different types of innovation outcomes at the country level (see Tab. 1). *Innov* denotes the percentage of firms that introduced a new product or service within the last three years, while *Innov_mark* represents the percentage of firms whose innovation was also new to the main market.

Tab. 1: Overview of variables – Part 1

Variables	Description	Theoretical justifications
Innov	Companies that introduced a new product/service over last 3 years	Incremental innovation (Natera & Castellacci, 2021; Teece, 2018)
Innov_mark	Companies whose new product/ service is also new in the main market	Radical innovation (Schilke et al., 2018; Teece, 2014)
Corrupt	Companies that choosing corruption as their biggest obstacle	Corruption distorts formal rules, increases transaction costs, and reduces innovation incentives (Bănică et al., 2024; North, 1990; Peng et al., 2020; Scott, 2014)
Labor_regul	Companies that identifying labor regulations as a major or very severe constraint	Regulatory burdens limit flexibility and increase costs, discouraging risk-taking and radical innovation (Gao & Anwar, 2024; Natera & Castellacci, 2021; Scott, 2014)

Tab. 1: Overview of variables – Part 2

Variables	Description	Theoretical justifications
Labor_prod	Real annual labor productivity growth (%)	Productivity reflects internal capabilities and organizational slack that enable innovation (Schilke et al., 2018; Teece, 2018)
Wom_top	Percentage of firms with a woman top manager	Gender diversity in top management enhances decision-making diversity, collaboration, and innovation capacity (Capozza & Divella, 2024; Jeong & Harrison, 2017; Lee & Chung, 2022)
Full_work	Number of permanent full-time equivalent workers	Firm size reflects resource endowment and organizational capacity influencing innovation (Beck et al., 2022; Leiponen & Helfat, 2010)
Train_work	Proportion of workers offered formal training over last fiscal year (%)	Training enhances absorptive capacity and knowledge diffusion, facilitating innovation (Hanushek & Woessmann, 2015)
Years_top	Years of the top manager's experience working in the firm's sector	Managerial experience improves strategic orientation and resource allocation, positively influencing innovation (Hitt et al., 2016)
Nation	Coarse contextual dummy (CEE = 1, otherwise 0); legacy label "Nation" retained for consistency; not a full set of country fixed effects	Contextual variable controlling for institutional and cultural heterogeneity across EU member states (Natera & Castellacci, 2021; Peng et al., 2020); accounts for shared institutional and historical legacies at a coarse regional level while avoiding overfitting in a small-*N* (27) design

Source: own based on WBES (2023)

Institutional factors include *Corrupt* (percentage of firms reporting corruption as the most severe obstacle) and *Labor_regul* (percentage identifying labor regulations as a major or very severe barrier). Firm-level capabilities are represented by additional country-level aggregates: *Labor_prod* (average annual growth of labor productivity), *Wom_top* (percentage of firms with a woman in top management), *Full_work* (average number of permanent employees per firm), *Train_work* (percentage providing formal training), and *Years_top* (average tenure of top managers in the sector). *Nation* identifies the country to account for contextual heterogeneity. The variable denotes a subregional dummy (CEE = 1). We keep the legacy label for consistency across the manuscript; it is not a full set of country fixed effects. A limitation of this approach is the small number of observations (27 countries), which restricts degrees of freedom and prevents capturing within-country heterogeneity. Results should therefore be interpreted as exploratory associations rather than causal effects. Nevertheless, WBES provides harmonized and nationally representative indicators, ensuring robust comparability and policy relevance.

2.3 Data analysis method

This study applies the ordinary least squares (OLS) method to estimate the relationship between two dependent variables: *Innov* (percentage of firms that introduced a new product or service within the last three years) and *Innov_mark* (percentage of firms whose innovation was also new to the main market) and a set of institutional and organizational predictors. OLS is a classical econometric technique that provides coefficient estimates by minimizing the sum of squared residuals, offering both transparency and interpretability in cross-sectional research. In line with the Gauss-Markov theorem, the assumptions of linearity, exogeneity, homoscedasticity, and absence of perfect multicollinearity were assessed. Heteroskedasticity-robust (White-consistent) standard errors were computed to address potential heteroskedasticity, and variance inflation factors (VIF) were used to test for multicollinearity. Given the relatively small sample size (27 countries), we favour parsimonious specifications and interpret estimates as betweencountry associations. Model fit was assessed through the coefficient of determination (R^2), F-tests of overall significance, and information criteria (AIC, Schwarz criterion, Hannan-Quinn). Two separate regression models were estimated. The model for *Innov_mark* explains 54% of the variance (R^2 = 0.541) and identifies four significant predictors: labor regulations, labor productivity growth, firm size, and gender diversity in top management. By contrast, the model for *Innov* explains 74% of the variance (R^2 = 0.739), but only labor productivity growth emerges as statistically significant, with a negative effect. These results highlight distinct determinants of incremental versus market-novel innovation, underscoring the importance of distinguishing innovation types in cross-country research. The use of OLS is motivated by its clarity, interpretability, and suitability for exploratory cross-country analysis, while also enabling the disentangling of marginal effects of institutional and organizational factors. For a comprehensive overview of linear regression methodology and its applications, see Montgomery et al. (2021). To avoid model saturation with N = 27, we do not estimate models with full country fixed effects. Instead, the *Nation* variable serves as a coarse subregional control (CEE = 1) to absorb broad contextual heterogeneity. We therefore favour parsimonious specifications and heteroskedasticity-robust standard errors. Given these constraints, cross-national analyses widen the scope of inference but can violate independence assumptions; results are thus interpreted as exploratory associations rather than causal effects (Claessens et al. 2023).

3 Results and discussion

The regression analysis revealed that the share of firms introducing products or services that are also new to the main market is significantly influenced by several key factors (see Tab. 2). Standard errors are heteroskedasticity-robust (White-consistent). *Nation* denotes a coarse subregional dummy (CEE = 1) used as a parsimonious context control, full country fixed effects are not included due to small-N (27). The strongest positive effect was observed for the number of permanent employees (β = 0.621; p = 0.012), suggesting that larger firms possess greater capacities for developing and implementing innovations with the potential to reach new market segments. Similarly, real labor productivity growth (β = 0.350; p = 0.018) emerged as a statistically significant determinant, confirming that higher efficiency in production and processes is associated with a greater likelihood of more radical innovations. Another noteworthy finding is the positive effect of having a woman in top management (β = 0.497; p = 0.027), which supports the literature emphasizing management diversity as a source of new perspectives and innovative thinking. The perception of labor regulations as a major constraint also shows a significant positive relationship (β = 0.363; p = 0.041), which can be interpreted as regulatory pressures paradoxically stimulating innovation, as firms search for new ways to remain competitive. Other predictors, such as the top manager's experience, the provision of employee training, or cultural factors, did not prove to be statistically significant. Overall, the model explains approximately 54% of the variability in the dependent variable (R^2 = 0.541; F(8, 18) = 2.66; p = 0.041), indicating satisfactory predictive power and underscoring the relevance of the identified determinants for understanding betweencountry differences in the national share of firms achieving market-novel innovation.

The regression results for the share of firms that introduced a new product or service within the last three years (see Tab. 3) indicate a somewhat different pattern of determinants compared to the introduction of products new to the main market. Among the examined predictors, two factors stand out as statistically significant. First, the perception of corruption as the biggest obstacle shows a negative and significant relationship (β = −0.472; p = 0.030), implying that higher perceived corruption in the business environment discourages firms from pursuing new product or service development. Second, real annual

Tab. 2: Regression analysis results for radical innovations

Independent	Coeff. estimate	Std. error	*t*-value	*p*-value	Sign. code
Const	−0.2381	0.2853	−0.8345	0.4149	
Labor_regul	0.3626	0.1642	2.2088	0.0404	**
Corrupt	−0.0455	0.1684	−0.2704	0.7899	
Labor_prod	0.3499	0.1338	2.6147	0.0176	**
Full_work	0.6210	0.2232	2.7825	0.0123	**
Train_work	−0.0568	0.1516	−0.3748	0.7122	
Years_top	0.2241	0.1625	1.3787	0.1849	
Wom_top	0.4974	0.2063	2.4117	0.0268	**
Nation	0.1645	0.1296	1.2696	0.2204	
R^2	0.541				

Note: * significant at significance level $p < 0.1$; ** significant at significance level $p < 0.05$; *** significant at significance level $p < 0$.

Source: own

labor productivity growth also exerts a significant negative effect (β = −0.364; $p = 0.006$), suggesting that productivity gains in this context are not associated with greater innovation output but may instead reflect efficiency-oriented strategies rather than product diversification. Other independent variables, including firm size, workforce training, managerial experience, gender diversity in top management, and cultural factors, did not reach statistical significance. The model explains approximately 74% of the variance in the dependent variable ($R^2 = 0.739$; $F(8, 18) = 4.40$; $p < 0.001$), which highlights its strong explanatory capacity. Overall, these findings suggest that systemic barriers such as corruption and structural shifts in productivity may significantly shape firms' innovation behavior, potentially limiting their engagement in the introduction of new products and services over time.

Tab. 3: Regression analysis results for incremental innovations

Independent	Coeff. estimate	Std. error	*t*-value	*p*-value	Sign. code
Const	0.5245	0.3400	1.5427	0.1403	
Labor_regul	−0.0458	0.1956	−0.2343	0.8174	
Corrupt	−0.4720	0.2007	−2.3525	0.0302	**
Labor_prod	−0.3643	0.1594	−2.2850	0.0347	**
Full_work	0.4078	0.2660	1.5333	0.1426	
Train_work	0.1035	0.1806	0.5731	0.5737	
Years_top	0.0109	0.1937	0.0560	0.9559	
Wom_top	0.0567	0.2458	0.2307	0.8202	
Nation	0.0920	0.1544	0.5957	0.5588	
R^2	0.739				

Note: * significant at significance level $p < 0.1$; ** significant at significance level $p < 0.05$; *** significant at significance level $p < 0$.

Source: own

The results reveal a clear divergence in how institutional and firm-level factors influence incremental versus market-novel innovation. Consistent with institutional theory, *H1a* is supported: perceived corruption significantly reduces incremental innovation. This aligns with the theoretical mechanisms outlined earlier, where corruption raises transaction costs, lowers appropriability, and distorts formal rules of the game (Bó & Rossi, 2007; De Rosa et al., 2015; North, 1990; Scott, 2014). These mechanisms particularly hinder routine, incremental innovation that depends on frequent regulatory interactions. In contrast, *H1b* is not confirmed, as corruption shows no significant effect on market novelty. This asymmetry is consistent with the theoretical expectation that radical, market-novel innovation relies more heavily on internal capabilities and strategic leadership (Schilke et al., 2018; Teece, 2014), making firms less vulnerable to institutional frictions highlighted in the theoretical section.

Regarding labor regulations, *H2a* is not supported. The absence of a significant effect on incremental innovation is consistent with the argument that incremental improvements rely on established routines and may be less sensitive to regulatory rigidity (Filippetti & Guy, 2020). However, *H2b* is rejected in an unexpected direction: labor regulations exhibit a positive relationship with market-novel innovation. Although this contradicts the conventional expectation that strict EPL restricts restructuring and risk-taking (Acharya et al., 2013; Bai et al., 2020), it is consistent with the compliance-induced innovation mechanism noted in the theoretical section, where regulatory stability and constraint may motivate firms to pursue more ambitious strategic differentiation (Bassanini et al., 2009; Gao & Anwar, 2024). This result enriches institutional theory by illustrating that regulatory pressure can stimulate, rather than solely inhibit, innovation under certain competitive conditions.

The findings for productivity display similar asymmetry. *H3b* is supported: higher labor productivity significantly increases the likelihood of market-novel innovation, confirming RBV arguments that productivity reflects internal capability, organizational slack, and absorptive capacity needed for high-risk projects (Barney, 1991; Schilke et al., 2018; Teece, 2018). In contrast, *H3a* shows a negative effect on incremental innovation, supporting the substitution mechanism described in the theoretical section, where efficiency gains reduce the need for small product updates (Hanushek & Woessmann, 2015; OECD, 2023). This extends the RBV perspective by demonstrating that productivity is not a uniform enhancer of innovation but has differentiated effects depending on innovation type.

Gender diversity reveals a similar pattern. *H4b* is supported: the presence of a woman in top management significantly increases the probability of market-novel innovation. This aligns with RBV insights and evidence that diverse leadership enhances knowledge integration, strategic adaptability, and openness to external ideas (Capozza & Divella, 2024; Jeong & Harrison, 2017; Lee & Chung, 2022). *H4a* is not supported, which is coherent with the idea that incremental innovation depends more on operational routines than on diversity-driven strategic capabilities.

Overall, the results confirm that institutional barriers, especially corruption, primarily constrain incremental innovation, while firm-level resources, particularly productivity, firm size, and leadership diversity are decisive drivers of market-novel innovation. This duality reflects the complementary logics of institutional theory and the RBV: institutional quality shapes the predictability and fairness of the external environment (North, 1990; Peng et al., 2020; Scott, 2014), whereas internal capabilities determine how effectively firms can transform resources into competitive advantage and ambitious innovations (Barney, 1991; Schilke et al., 2018; Teece, 2014). The asymmetric effects identified for labor regulation (*H2b*) and productivity (*H3a*) further refine both frameworks by showing that institutional pressures and organizational capabilities affect different innovation types through distinct mechanisms.

Conclusions

This study examined how institutional and organizational factors (perceived corruption, labor regulations, productivity growth, firm size and gender diversity) shape incremental and market-novel innovation across 27 EU countries. Our findings challenge the conventional assumption that institutional barriers uniformly hinder innovation and show that their impact depends on innovation type and firm-level resources. Corruption significantly suppresses incremental innovation, consistent with institutional theory linking bribery to higher transaction costs and lower appropriability (North, 1990; Scott, 2014). In contrast, corruption does not affect market-novel innovation, indicating that radical innovation relies more on internal strategic capabilities and long-term resources (Schilke et al., 2018; Teece, 2014). Labor regulations display a similar asymmetry: while unrelated to incremental innovation, they positively predict market-novel outcomes, aligning with compliance-induced innovation mechanisms. Productivity growth also reinforces this duality, reducing incremental innovation while enabling market novelty through greater resource slack and absorptive capacity (Hanushek & Woessmann, 2015; OECD, 2023). Firm size and gender diversity further strengthen firms' ability to pursue market-novel innovation, supporting RBV views that organizational capacity and inclusive leadership enhance strategic adaptability (Capozza & Divella, 2024; Jeong & Harrison, 2017). Overall, incremental innovation is shaped mainly by institutional weaknesses, whereas market-novel innovation is driven by firms' internal resource endowments. These findings refine institutional theory and the RBV by showing that institutional pressures and internal capabilities operate through innovation-type-specific mechanisms, offering a multi-level explanation of innovation performance in the EU.

Practical and theoretical implications. The findings carry practical and theoretical implications. For managers, they underline the need to strengthen governance and compliance in high-corruption environments, as corruption weakens innovation incentives (Cuervo-Cazurra, 2016). Labor regulations can provide stability when paired with long-term workforce development, enabling firms to convert institutional constraints into strategic advantages that support market-novel innovation (Acharya et al., 2013). Productivity-enhancing initiatives, such as upskilling, digitalization and efficient resource allocation, remain essential for creating the slack required for ambitious innovation (Hanushek & Woessmann, 2015; OECD, 2023). Gender-diverse top management further strengthens adaptability and knowledge integration (Capozza & Divella, 2024; Jeong & Harrison, 2017). For EU policymakers, the findings recommend combining support for innovation with strengthening integrity and enforcement in areas with high levels of discretion, such as complex electronic public procurement, digital licensing, and risk-based controls, as perceived barriers often reflect actual enforcement frictions that dampen innovation outcomes (Hallward-Driemeier & Pritchett, 2015; The World Bank, 2019–2023).

Theoretically, the study shows that institutional theory explains why corruption constrains incremental innovation by increasing uncertainty and undermining routinized innovation (De Rosa et al., 2015). In contrast, the RBV accounts for why productivity, firm size and leadership diversity support market-novel outcomes through enhanced absorptive capacity (Beck et al., 2022; Jeong & Harrison, 2017). The positive effect of labor regulation on market novelty nuances debates by showing that institutional pressures may also stimulate experimentation (Acharya et al., 2013). Overall, this framework integrates institutional constraints with firm-level capabilities and explains why similar institutional conditions can lead firms to different outcomes in innovation (Barney, 1991; Hallward-Driemeier & Pritchett, 2015; Teece, 2014).

Limitations and future research. The study's cross-sectional design, perception-based measures and limited number of country-level observations restrict causal interpretation. Future research should explore micro-level mechanisms behind compliance-induced innovation, including managerial practices and workforce strategies that convert regulatory constraints into innovation opportunities. Comparative research in non-EU or emerging-economy contexts could test whether similarly asymmetric patterns, especially the non-effect of corruption on radical innovation, generalize beyond Europe. Longitudinal datasets would further allow deeper analysis of how institutional reforms, productivity dynamics and leadership diversity jointly shape innovation trajectories.

Acknowledgments: *This paper was supported by the Student Grant Competition of the University of Pardubice SGS_2025_018. We thank the Enterprise Analysis Unit of the Development Economics Global Indicators Department of the World Bank for the data.*

References

Acharya, V. V., Baghai, R. P., & Subramanian, K. V. (2013). Labor laws and innovation. *The Journal of Law and Economics*, *56*(4), 997–1037. https://doi.org/10.1086/674106

Alfano, V., Capasso, S., & Goel, R. K. (2021). EU accession: A boon or bane for corruption? *Journal of Economics and Finance*, *45*(1), 1–21. https://doi.org/10.1007/s12197-020-09522-8

Bai, J., Fairhurst, D., & Serfling, M. (2020). Employment protection, investment, and firm growth. *The Review of Financial Studies*, *33*(2), 644–688. https://doi.org/10.1093/rfs/hhz066

Bănică, A., Ţigănaşu, R., Nijkamp, P., & Kourtit, K. (2024). Institutional quality in green and digital transition of EU regions – A recovery and resilience analysis. *Global Challenges*, *8*(9), 2400031. https://doi.org/10.1002/gch2.202400031

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, *17*(1), 99–120. https://doi.org/10.1177/014920639101700108

Bartelsman, E., Haltiwanger, J., & Scarpetta, S. (2013). Cross-country differences in productivity: The role of allocation and selection. *American Economic Review*, *103*(1), 305–334. https://doi.org/10.1257/aer.103.1.305

Bassanini, A., Nunziata, L., & Venn, D. (2009). Job protection legislation and productivity growth in OECD countries. *Economic Policy*, *24*(58), 349–402. https://doi.org/10.1111/j.1468-0327.2009.00221.x

Beck, S., Bergenholtz, C., Bogers, M., Brasseur, T.-M., Conradsen, M. L., Di Marco, D., Distel, A. P., Dobusch, L., Dörler, D., Effert, A., Fecher, B., Filiou, D., Frederiksen, L., Gillier, T., Grimpe, C., Gruber, M., Haeussler, C., Heigl, F., Hoisl, K., ..., Xu, S. M. (2022). The open innovation in science research field: A collaborative conceptualisation approach. *Industry and Innovation*, *29*(2), 136–185. https://doi.org/10.1080/13662716.2020.1792274

Boly, V., Enjolras, M., Husson, S., Morel, L., Dupont, L., & Benis, L. (2022). Innovation capacity of city administrations: A best practices approach. *Journal of Innovation Economics & Management*, *38*(2), 169–198. https://doi.org/10.3917/jie.038.0169

Capozza, C., & Divella, M. (2024). Gender diversity in European firms and the R&D-innovation-productivity nexus. *The Journal of Technology Transfer*, *49*(3), 801–822. https://doi.org/10.1007/s10961-023-10003-3

Claessens, S., Kyritsis, T., & Atkinson, Q. D. (2023). Cross-national analyses require additional controls to account for the non-independence of nations. *Nature Communications*, *14*(1), 5776. https://doi.org/10.1038/s41467-023-41486-1

Cuervo-Cazurra, A. (2016). Corruption in international business. *Journal of World Business*, *51*(1), 35–49. https://doi.org/10.1016/j.jwb.2015.08.015

Dal Bó, E., & Rossi, M. A. (2007). Corruption and inefficiency: Theory and evidence from electric utilities. *Journal of Public Economics*, *91*(5–6), 939–962. https://doi.org/10.1016/j.jpubeco.2006.11.005

De Rosa, D., Gooroochurn, N., & Görg, H. (2015). Corruption and productivity: Firm-level evidence. *Jahrbücher für Nationalökonomie und Statistik, 235*(2), 115–138.

Dezsö, C. L., & Ross, D. G. (2012). Does female representation in top management improve firm performance? A panel data investigation. *Strategic Management Journal*, *33*(9), 1072–1089. https://doi.org/10.1002/smj.1955

Dokas, I., Panagiotidis, M., Papadamou, S., & Spyromitros, E. (2023). Does innovation affect the impact of corruption on economic growth? International evidence. *Economic Analysis and Policy*, *77*, 1030–1054. https://doi.org/10.1016/j.eap.2022.12.032

Donges, A., Meier, J.-M., & Silva, R. C. (2023). The impact of institutions on innovation. *Management Science*, *69*(4), 1951–1974. https://doi.org/10.1287/mnsc.2022.4403

Ellis, J., Smith, J., & White, R. (2020). Corruption and corporate innovation. *Journal of Financial and Quantitative Analysis*, *55*(7), 2100–2130. https://doi.org/10.1017/S0022109019000735

Filippetti, A., & Guy, F. (2020). Labor market regulation, the diversity of knowledge and skill, and national innovation performance. *Research Policy*, *49*(1), 103867. https://doi.org/10.1016/j.respol.2019.103867

Fisman, R., Guriev, S., Ioramashvili, C., & Plekhanov, A. (2024). Corruption and firm growth: Evidence from around the world. *The Economic Journal*, *134*(660), 1494–1516. https://doi.org/10.1093/ej/uead100

Gao, L., & Anwar, S. (2024). Labour market rigidity and firm innovation. *Australian Economic Papers*, *63*(2), 237–257. https://doi.org/10.1111/1467-8454.12326

Hall, B. H. (2011). *Innovation and productivity* [working paper No. w17178]. National Bureau of Economic Research. https://doi.org/10.3386/w17178

Hall, B. H., Lotti, F., & Mairesse, J. (2009). Innovation and productivity in SMEs: Empirical evidence for Italy. *Small Business Economics*, *33*(1), 13–33. https://doi.org/10.1007/s11187-009-9184-8

Hallward-Driemeier, M., & Pritchett, L. (2015). How business is done in the developing world: Deals versus rules. *Journal of Economic Perspectives*, *29*(3), 121–140. https://doi.org/10.1257/jep.29.3.121

Hanushek, E. A., & Woessmann, L. (2015). *The knowledge capital of nations: Education and the economics of growth*. The MIT Press. https://doi.org/10.7551/mitpress/9780262029179.001.0001

Hitt, M. A., Ireland, R. D., & Hoskisson, R. E. (2016). *Strategic management: Concepts and cases: Competitiveness and globalization*. Cengage Learning.

Jeong, S.-H., & Harrison, D. A. (2017). Glass breaking, strategy making, and value creating: Meta-analytic outcomes of women as CEOs and TMT members. *Academy of Management Journal*, *60*(4), 1219–1252. https://doi.org/10.5465/amj.2014.0716

Lee, C.-C., Wang, C.-W., & Ho, S.-J. (2020). Country governance, corruption, and the likelihood of firms' innovation. *Economic Modelling*, *92*, 326–338. https://doi.org/10.1016/j.econmod.2020.01.013

Lee, J., & Chung, J. (2022). Women in top management teams and their impact on innovation. *Technological Forecasting and Social Change*, *183*, 121883. https://doi.org/10.1016/j.techfore.2022.121883

Leiponen, A., & Helfat, C. E. (2010). Innovation objectives, knowledge sources, and the benefits of breadth. *Strategic Management Journal*, *31*(2), 224–236. https://doi.org/10.1002/smj.807

Mironov, M. (2015). Should one hire a corrupt CEO in a corrupt country? *Journal of Financial Economics*, *117*(1), 29–42. https://doi.org/10.1016/j.jfineco.2014.03.002

Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to linear regression analysis*. John Wiley & Sons.

Moretti, E. (2021). The effect of high-tech clusters on the productivity of top inventors. *American Economic Review*, *111*(10), 3328–3375. https://doi.org/10.1257/aer.20191277

Natera, J. M., & Castellacci, F. (2021). Transformational complexity, systemic complexity and economic development. *Research Policy*, *50*(7), 104275. https://doi.org/10.1016/j.respol.2021.104275

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press. https://doi.org/10.1017/CBO9780511808678

OECD. (2023). *OECD science, technology and industry scoreboard 2023: Innovation and growth*. OECD Publishing.

Peng, L., Li, Y., van Essen, M., & Peng, M. W. (2020). Institutions, resources, and strategic orientations: A meta-analysis. *Asia Pacific Journal of Management*, *37*(2), 499–529. https://doi.org/10.1007/s10490-018-09642-0

Quan, X., Zhang, K., Zhong, R., & Zhu, Y. (2023). Political corruption and green innovation. *Pacific-Basin Finance Journal*, *82*, 102169. https://doi.org/10.1016/j.pacfin.2023.102169

Quintini, G. (2011). Over-qualified or under-skilled: A review of existing literature [OECD social, employment and migration working paper No. 121]. In *OECD social, employment and migration working papers*. Organisation for Economic Co-Operation and Development (OECD). https://doi.org/10.1787/5kg58j9d7b6d-en

Schilke, O., Hu, S., & Helfat, C. E. (2018). Quo vadis, dynamic capabilities? A content-analytic review of the current state of knowledge and recommendations for future research. *Academy of Management Annals*, *12*(1), 390–439. https://doi.org/10.5465/annals.2016.0014

Scott, W. R. (2014). *Institutions and organizations: Ideas, interests, and identities* (4th ed.). Sage Publications.

Shabir, M. I., Rehman, H. U., & Khan, R. E. A. (2020). Effect of corruption on firm level innovation: Evidence from Pakistan. *Economics and Business Letters*, *9*(1), 41–47. https://doi.org/10.17811/ebl.9.1.2020.41-47

Soukupova, N., Dvorakova, A., & Kotkova Striteska, M. (2024). Investigating the influence of gender diversity and management practices on the introduction of innovation in CEE countries. *Proceedings of European Conference on Management, Leadership & Governance. Academic Conferences International Limited, 20*(1), 562–569. https://doi.org/10.34190/ecmlg.20.1.3037

Stříteská, M., Zapletal, D., & Jelínková, L. (2016). Performance management systems in Czech companies: Findings from a questionnaire survey. *E&M Economics and Management*, *19*(4), 44–55. https://doi.org/10.15240/tul/001/2016-4-004

Tang, S., Nadkarni, S., Wei, L., & Zhang, S. X. (2021). Balancing the yin and yang: TMT gender diversity, psychological safety, and firm ambidextrous strategic orientation in Chinese high-tech SMEs. *Academy of Management Journal*, *64*(5), 1578–1604. https://doi.org/10.5465/amj.2019.0378

Teece, D. J. (2014). The foundations of enterprise performance: Dynamic and ordinary capabilities in an (economic) theory of firms. *Academy of Management Perspectives*, *28*(4), 328–352. https://doi.org/10.5465/amp.2013.0116

Teece, D. J. (2018). Business models and dynamic capabilities. *Long Range Planning*, *51*(1), 40–49. https://doi.org/10.1016/j.lrp.2017.06.007

The World Bank. (2019). *World Bank enterprise surveys*. https://www.enterprisesurveys.org

Van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology*, *89*(6), 1008–1022. https://doi.org/10.1037/0021-9010.89.6.1008

Yakovlev, E., & Zhuravskaya, E. (2013). The unequal enforcement of liberalization: Evidence from Russia's reform of business regulation. *Journal of the European Economic Association*, *11*(4), 808–838. https://doi.org/10.1111/jeea.12026

Zhong, W., & Zheng, Y. (2025). Do political connections stifle firm innovation? Natural experimental evidence from China's anti-corruption campaigns. *Strategic Management Journal*, *46*(5), 1183–1216. https://doi.org/10.1002/smj.3690

Using an AHP-GIS based framework to develop a resilient and competitive logistics network: A study of container yard location covering multimodal railway links in Northeastern Thailand

Prin Weerapong[1], Jakub Dyntar[2]

[1] *Rajamangala University of Technology Thanyaburi, Faculty of Business Administration, Department of Logistics and Supply Chain Management, Thailand, ORCID: 0009-0009-5229-6185, prin_w@rmutt.ac.th;*

[2] *Technical University of Liberec, Faculty of Economics, Department of Business Administration and Management, Czech Republic, ORCID: 0000-0002-5927-9443, jakub.dyntar@tul.cz (corresponding author).*

Abstract: *This study examines the optimal location for container yard development in Northeastern Thailand to enhance multimodal transport connectivity and strengthen national competitiveness in regional and global trade. Utilizing the analytic hierarchy process (AHP) and geographic information system (GIS), the study evaluated 18 railway stations based on five key factors: logistics suitability (13.5%), infrastructure readiness (14.6%), government support (22.5%), railway network capabilities (16.9%), and investment economics (32.5%). The findings identify Nata Station in Nong Khai Province as the most strategically advantageous location, with the highest composite weighted Z-score of 0.8775. The station features a road network density of 0.85 km per square kilometer and a railway density of 0.32 km per square kilometer, facilitating efficient cross border freight transportation. The establishment of a container yard in this optimal location is expected to reduce logistics costs, improve supply chain efficiency, and enhance cross border trade, particularly with member states of the Association of South East Asian Nations (ASEAN) trading bloc and China. This development reinforces the role of Thailand as a regional logistics hub, fostering industrial expansion and economic integration. Furthermore, government backed policies, including investment incentives and infrastructure development, contribute to building a resilient and competitive logistics network. The study provides a strategic framework for policymakers and private sector stakeholders to leverage logistics infrastructure in strengthening the long-term competitiveness of Thailand in global trade.*

Keywords: *Supply chain management, multi criteria decision analysis, analytic hierarchy process, geographic information system, site selection, container yard.*

JEL Classification: *R42, C61, O21, F14.*

APA Style Citation: Weerapong, P., & Dyntar, J. (2026). Using an AHP-GIS based framework to develop a resilient and competitive logistics network: A study of container yard location covering multimodal railway links in Northeastern Thailand. *E&M Economics and Management, 29*(2), 156–176. https://doi.org/10.15240/tul/001/2026-2-011

Introduction

The development of logistics and transportation infrastructure plays a crucial role in driving regional and national economic growth, particularly in an era of continuously expanding international trade and transportation. An efficient logistics system is essential for enhancing national competitive advantage in the global marketplace (Cahyono et al., 2023). The development of logistics infrastructure continues to face various limitations and challenges that need to be systematically addressed through comprehensive research and planning. In recent years, multiple government agencies in Thailand have conducted studies and developed plans to enhance the Thai logistics system, e.g., Garrow et al. (2021) or Theerathitichaipa et al. (2024). As these studies have predominantly concentrated on international trade aspects, they unfortunately lack detailed analysis of domestic transportation and regional considerations (Puchongkawarin & Ransikarbum, 2021). Furthermore, there exists inadequate integration between policy perspectives and practical implementation considerations (Prasittisopin et al., 2024), as well as limited consideration of environmental and social impacts in infrastructure development planning (Bui & Nguyen, 2022). This gap in research highlights the need for more comprehensive studies that address both international and domestic logistics requirements.

The northeastern region of Thailand presents significant potential for development as a logistics hub due to its strategic location, which enables transportation linkages both domestically and internationally, particularly with neighboring countries in the Indochina region (Dai, 2022). Yet the development of container yards in this region lacks a systematic and comprehensive study on an optimal location selection. This study aims to identify suitable areas for container yard development along railway routes to enhance logistics management efficiency at provincial, regional, and international levels. Thus, we suggest a framework using complex quantitative analysis tools and integration of quantitative and qualitative data to seek and prioritize container yard locations and associated logistics infrastructure to support economic expansion, distribute development opportunities, and improve quality of life in provincial and regional areas. Within the proposed framework we employ a mixed method research approach, applying the analytic hierarchy process (AHP) to analyze and prioritize suitable areas. The analysis considers logistics advantages, infrastructure readiness, government support, railway network capabilities, and economic feasibility as key determining factors (Wangai et al., 2020). The research methodology has been designed to ensure comprehensive evaluation of all relevant aspects while maintaining analytical rigor. It employs a comprehensive approach to location analysis that incorporates multiple stakeholder perspectives and robust analytical methods. This ensures that the selected sites maximize economic benefits while minimizing potential negative impacts on local communities and the environment (Abu Aisha et al., 2020). The methodology integrates both quantitative metrics and qualitative assessments to provide a holistic evaluation framework for site selection (Khorram, 2020).

The rest of this paper in organized as follows: in the first section, criteria for container yard location are identified using a review of relevant scientific literature; in the second section, a framework for efficient site selection is introduced including required data sources; then in the third section, the outputs are presented for the Northeastern Thailand region and discussed in detail. Finally, in the last section, the paper is concluded.

1 Theoretical background

Multimodal transportation has emerged as a fundamental driving force of economic development in the modern global economy. According to Kramarz and Przybylska (2021), multimodal transport, defined as the movement of goods using two or more modes of transportation under a single contract, significantly enhances supply chain efficiency. This efficiency gain directly contributes to economic development through reduced logistics costs and improved market access. Bandyopadhyay and Bhatnagar (2023) further establish that multimodal transportation systems generate substantial economies of scale by documenting a reduction in total logistics costs across developing economies. The implementation of these systems, as noted by Jiao et al. (2020), creates significant multiplier effects throughout regional economies. Within multimodal transportation systems, selecting the optimal position of logistics hubs using location theory represents a critical task. Location theory has evolved significantly in its application to competitive trade advantage, with contemporary research emphasizing the dynamic interplay between spatial economics and market competitiveness. Li et al. (2020) assert that optimal location selection serves as a fundamental determinant of competitive advantage, particularly in logistics intensive industries. Building on this foundation, Aldrighetti et al. (2021) demonstrate that strategic location decisions can reduce operational costs while enhancing market responsiveness. Herzog (2021) expand this perspective by identifying key location factors, including transportation costs, market accessibility, and labor availability, as critical determinants of competitive success. Their research reveals that firms with optimally selected locations achieve higher operational efficiency compared to their competitors. Zhou et al. (2021) emphasize the importance of supply chain considerations in location decisions, demonstrating that well located facilities can reduce total logistics costs. Correia and Melo (2021) further establish that optimal location selection significantly influences customer response times and operational flexibility, with strategically located facilities showing faster market response rates compared to poorly situated competitors.

The selection of optimal locations for container yards represents a critical factor in developing successful logistics networks. Recent research by Kaliszewski et al. (2020) demonstrates that logistics attractiveness significantly influences location potential, with well positioned container yards showing higher operational efficiency. Danyluk (2019) identifies key success factors, including network connectivity and infrastructure readiness, emphasizing the importance of accessibility to raw materials and markets. Moskvichev et al. (2021) reveal that strategically located container yards reduce transportation costs compared to suboptimal locations. De Langen et al. (2020) highlight the significance of business ecosystem development, emphasizing that regions with established container yard clusters attract more related businesses. Song (2021) demonstrates that efficient container storage and distribution capabilities can improve supply chain performance. Zhou et al. (2021) conclude that successful container yard locations require careful consideration of both immediate operational needs and long-term development potential, with their research showing that comprehensive location analysis leads to improved long term sustainability metrics. When studying the relevant literature, 5 main criteria for container yard location selection were identified each consisting of a certain set of sub criteria suitable to support decision making about container yard location (see Tab. 1).

The relevant literature used in Tab. 1 came from a structured review conducted using the Scopus and Web of Science databases. The search strategy combined keywords related to logistics facility location and multi-criteria decision-making. The primary search query was (“container yard” OR “dry port” OR “logistics hub” OR “inland terminal”) AND (“location selection” OR “site selection”) AND (“AHP” OR “MCDA” OR “multi-criteria decision making”) AND (“logistics” OR “transportation”). The review covered the period 2019–2024, focusing on recent developments in logistics infrastructure and decision-support methods. The selection process followed three stages encompassing an initial retrieval returning approximately 120–150 promising articles, then a screening reducing these articles to 60 that finally became the subject of full-text analysis.

Besides quite common criteria such as logistics suitability or investment economics that are well known, for example from the local distribution logistics, this kind of complex decision-making problem also requires consideration of government participation or compliance with standards for reducing the environmental impact from noise, air and visual pollution. For instance, Göçer et al. (2022) demonstrate that regions with strong government support in supply chain management policy develop higher international competitiveness ratings. Xiong et al. (2020) reveal that public-private partnerships backed by government initiatives achieve better operational efficiency compared to purely private ventures. Curry et al. (2021) demonstrate that government supported supply chain management innovation programs result in higher technology adoption rates. Lo Storto and Evangelista (2023) identify government budget allocation as critical for large scale infrastructure development. Similarly, Jiang et al. (2021) emphasize environmental impact considerations, showing that railway transportation reduces carbon emissions compared to road transport. Khan et al. (2022) highlight the importance of sustainable infrastructure development, showing that green logistics

Tab. 1: Identified criteria for container yard location selection

Main criterion *i*	Sub criterion *i.j*	Source
1. Logistics suitability	1.1 Logistics operations attractiveness	Fahim et al., 2022; Sakai et al., 2020
	1.2 Number of logistics business operations	
	1.3 Storage and distribution efficiency	
2. Infrastructure	2.1 Labor infrastructure	Fahim et al., 2022; Tadić et al., 2020
	2.2 Cargo transportation infrastructure	
	2.3 Sustainable infrastructure development	
3. Government support	3.1 Logistics system development policy	Kaliszewski et al., 2020; Liu et al., 2021
	3.2 Investment support budget	
4. Railway network	4.1 Future railway network connectivity potential	Fahim et al., 2022; Khaslavskaya & Roso, 2020; Tadić et al., 2020; Ulutaş et al., 2020
	4.2 Railway development capability	
	4.3 Existing equipment utilization capability	
	4.4 Existing railway network accessibility	
	4.5 Utility infrastructure potential of candidate and adjacent areas	
	4.6 Environmental impact (noise/air/visual pollution)	
5. Investment economics	5.1 Land acquisition feasibility	Kaliszewski et al., 2020; Tadić et al., 2020
	5.2 Future development plan integration	
	5.3 Investment return potential	
	5.4 Regional value creation potential	

Source: own processing based on a systematic literature review

initiatives result in lower operational costs while reducing environmental impact.

Based on the nature of the decision-making problem, multi criteria decision analysis (MCDA) has emerged as an efficient tool in optimizing container yard location decisions. Belay et al. (2022) emphasize the effectiveness of MCDA in evaluating complex factors affecting infrastructure placement, while Pazzini et al. (2022) demonstrate its success in balancing quantitative metrics with qualitative considerations. Baydaş et al. (2024) show that MCDA based decisions result in higher operational efficiency compared to traditional single criterion approaches. According to Yannis et al. (2020), successful MCDA implementation requires careful consideration of both operational metrics and environmental impacts, with their study revealing improvement in decision quality. Furthermore, Abouhawwash and Jameel (2023) demonstrate the effectiveness of MCDA in evaluating financial feasibility. Within MCDA, the AHP serves as a comprehensive multi criteria decision making methodology for complex problem analysis. Kayikci (2021) demonstrate the effectiveness of following the AHP in logistics infrastructure planning, while Tavana et al. (2023) highlight its success in reducing decision bias. Despite extensive research on multimodal transportation and the selection of container yard locations, substantial deficiencies persist in addressing the integration of dynamic economic, technological, and policy-related factors. Current models predominantly depend on static decision-making frameworks, neglecting time-sensitive variables such as variations in land values, progress in transportation technology, and changing government rules that affect long-term viability. Furthermore, there is a paucity of research on the interaction between infrastructure expansion and state-sponsored incentives, which are essential for the smooth operation of multimodal logistics networks. Although elements like digital infrastructure, automation in freight movement, and environmental sustainability are gaining prominence in logistics research, there is an absence of recognized criteria for assessing their long-term effects on container yard site selection. A comprehensive strategy that incorporates spatial analytics, economic forecasting, and adaptive modelling methodologies is crucial to accurately represent the complexities of multimodal freight operations in Northeastern Thailand.

2 Research methodology

The goal of this research is to locate the most suitable site for container yard area development serving multimodal railway links in Northeastern Thailand. The study area contains 18 possible strategic railway stations (see Tab. 2) spread in the territory as shown in Fig. 1.

The research used a comprehensive methodological framework incorporating criteria for container yard location selection derived from analysis of the relevant literature (i.e., Tab. 1), AHP execution to assess the importance of criteria, *Z*-score modeling to enhance location optimization analysis, and geographic information system (GIS) analysis to further ensure the comprehensive evaluation of promising locations.

In the AHP methodology, three standardized key steps were followed including:

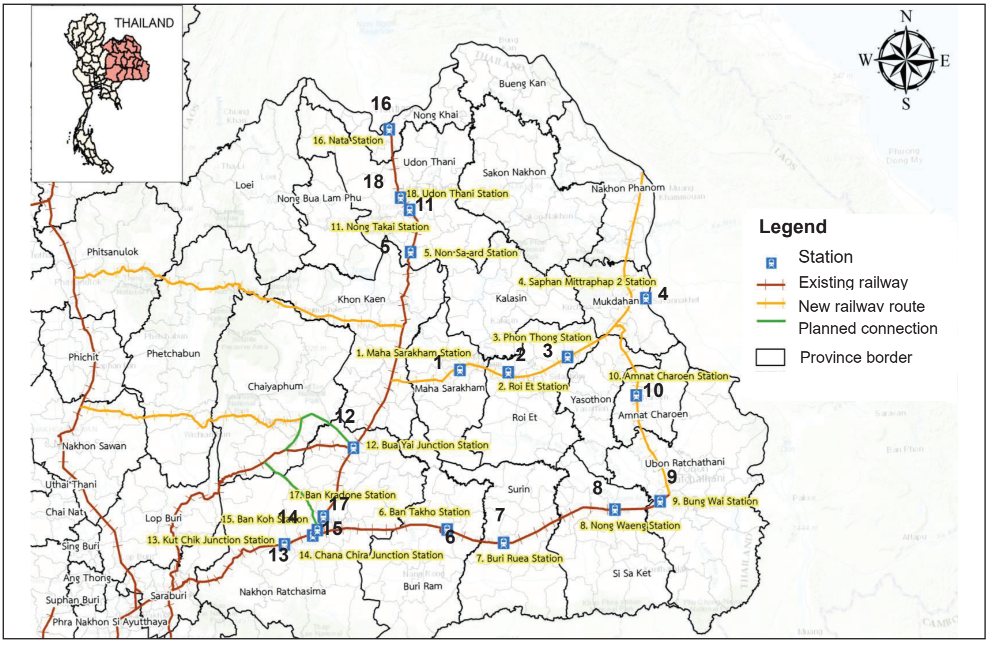

Fig. 1: **Study area map of railway stations in Northeastern Thailand with potential container yard development sites**

Source: own processing based on data from Department of Rail Transport of Thailand

Tab. 2: Railway stations for the location of the container yard area

Station *k*	Station name	Province
1	Maha Sarakham	Maha Sarakham
2	Roi Et	Roi Et
3	Phon Thong	Roi Et
4	Saphan Mittraphap	Mukdahan
5	Non-Sa-ard	Udon Thani
6	Ban Takho	Buriram
7	Buri Ruea	Surin
8	Nong Waeng	Sisaket
9	Bung Wai	Ubon Ratchathani
10	Amnat Charoen	Amnat Charoen
11	Nong Takai	Udon Thani
12	Bua Yai Junction	Nakhon Ratchasima
13	Kut Chik Junction	Nakhon Ratchasima
14	Chana Chira Junction	Nakhon Ratchasima
15	Ban Koh	Nakhon Ratchasima
16	Nata	Nong Khai
17	Ban Kradone	Nakhon Ratchasima
18	Udon Thani	Udon Thani

Source: own

Step 1: Hierarchy construction and AHP data collection

The decision problem was structured into a hierarchy, starting with the goal at the top (i.e., to select the optimal location for a container yard in Northeastern Thailand), followed by the main and sub criteria derived from the literature review (see Tab. 1) then evaluation of alternative sites at the bottom level (see Tab. 2). Pairwise comparisons of main and sub criteria were performed with respect to their importance to elements at higher level. These comparisons use Saaty's fundamental scale ranging from 1 to 9, where 1 indicates equal importance, 3 indicates moderate importance, 5 indicates strong importance, 7 indicates very strong importance, 9 indicates extreme importance, and 2, 4, 6, 8 represent intermediate values. The principal primary AHP data collection comprised stakeholder interviews and responses to an AHP questionnaire from three essential groups including government agencies, relevant transportation associations, and logistics related business organizations. Experts were selected using purposive sampling to ensure both expertise and representation across key stakeholder groups. Selection criteria included seniority, institutional relevance, and having a direct involvement in logistics or infrastructure planning. More specifically, seniority refers to positions with decision-making authority (e.g., directors, senior managers, committee members); institutional relevance refers to organizations directly involved in logistics, transport infrastructure, or related policy domains; and direct involvement refers to practical experience in logistics operations, infrastructure planning, or policy implementation. These criteria were evaluated using publicly available professional information and organizational roles rather than a formal quantitative scoring system. Initially, a broader pool of 45 candidates was identified, and contacted, of which 20 agreed to participate and complete the full pairwise comparison process, resulting in a response rate of approximately 44%. This level of participation is considered adequate for AHP-based studies, where the quality of expert judgment is prioritized over large sample sizes (Xu et al., 2022). The affiliations and roles of experts participating in AHP data collection are described in Tab. 3.

The interviews were performed through online meetings to accommodate senior executives from pertinent firms, while quantitative data was gathered using Google Forms, where experts assigned grades based on defined criteria. The data gathering period extended from October 2023 to May 2024.

Step 2: Calculation of priority vector and relative weights of criteria

Based on elements a_{ij} in a pairwise comparison matrix A (see Equation (1)), a priority vector (c_{ij})

Tab. 3: Affiliations and roles of experts participating in AHP data collection

Person	Affiliation	Role
1	Thai Chamber of Commerce and Board of Trade of Thailand	Vice Chairman of the Logistics and Supply Chain Committee
2		Senior Economic Structure Officer
3	The Federation of Thai Industries	Members of the Logistics and Supply Chain Subcommittee
4		
5	The Thai National Shippers' Council	Vice President of the Thai National Shippers' Council
6		Logistics Analyst
7	Import and Export Freight Transport Association	President of the Import and Export Freight Transport Association
8		Vice President of the Import and Export Freight Transport Association
9		Deputy President of the Thai Transport and Logistics Association
10	Trade Associations for Agricultural and Industrial Products	Vice President of the Thai Rice Exporters Association
11		Board Member of the Thai Rubber Association
12		President of the Thai Pulp and Paper Industries Association
13		President of the Thai Tapioca Products Factory Association
14	Thai Airfreight Forwarders Association (TAFA)	Vice President of TAFA
15	Port and Industrial Estate Authorities	Director of the Planning Division at Laem Chabang Port
16		Assistant Director of Technical Affairs at Map Ta Phut Industrial Port
17		Director of Songkhla and Phuket Deep Sea Ports
18		Head of Container Operations at Songkhla Deep Sea Port
19		Vice Chairman of the Logistics & Supply Chain Committee of Udon Thani Industrial Estate
20		Freight Operations Officer Level 10, Single Rail Transfer Operator Project, Laem Chabang Port

Source: own

is calculated using Equation (2) and the relative weight of each criterion (w_i) is calculated using Equation (3):

$$A = \begin{bmatrix} 1 & a_{12} & \dots & a_{1n} \\ 1/a_{12} & 1 & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n} & 1/a_{2n} & \dots & 1 \end{bmatrix} \quad (1)$$

$$c_{ij} = \frac{a_{ij}}{\sum_{i=1}^{n} a_{ij}} \quad (2)$$

$$w_i = \frac{\sum_{j=1}^{n} c_{ij}}{n} \quad (3)$$

Step 3: Consistency check

The consistency of judgments was verified using the consistency ratio (*CR*) based on Equation (4):

$$CR = \frac{CI}{RI} \quad (4)$$

where: *RI* represents the randomness index and *CI* represents the consistency index that can be calculated using Equation (5):

$$CI = \frac{\lambda_{Max} - n}{n - 1} \quad (5)$$

where: λ_{max} represents principal eigenvalue and *n* represents the pairwise comparison matrix size. *RI* values are summarized in Tab. 4.

A consistency ratio value ≤ 0.1 indicates acceptable consistency in the judgments.

Based on Bhuvaneskumar et al. (2023), in the *Z*-score the modelling steps are followed, including:

Step 1:

Z-score calculation ($Z_{k,j}$) for a k^{th} station and a j^{th} criterion using Equation (6):

$$Z_{k,j} = \frac{X_{k,j} - \mu_j}{\sigma_j} \quad (6)$$

where: $X_{k,j}$ represents raw value of station *k* for criterion *j*; μ_j represents mean of the values of all stations for criterion *j;* and σ_j represents standard deviation of the values for criterion *j*.

Sub criteria were derived from secondary data analysis of several key national policies and strategic plans. These include the National Strategic Plan (2018–2037) Infrastructure, Logistics and Digital Systems Master Plan (Issue 7); the Thai Logistics System Development Strategy (Phase 3, 2017–2021); a Government Policy Statement to Parliament (July 25, 2019); the 50-Year National Spatial Development Policy Plan (2057) with its 5, 10, and 15-year strategic roadmaps; the 20-Year Strategic Plan of the State Railway of Thailand for Economic Corridors Development under the Greater Mekong Sub region; and the Economic Cooperation Program and the Pan Asian Railway Network (Singapore-Kunming Rail Link). This research also incorporates insights from historical development studies, including various OTP reports such as the 2006 Multimodal Transportation and Logistics Management System Development Study, the 2006 Thai Cross-border Cargo Transportation Potential Development Study, the 2012 Logistics Value Chain Development Study for North-South and East-West Economic Corridors, the 2013 Multimodal Transportation Efficiency Enhancement Study, the Spur Lines Study issued by the State Railway of Thailand, and the 2018 Dry Port Development Master Plan Study, along with the 2017–2022 budget plan. Data sources used to evaluate each sub criterion are summarized in Tab. 5.

Tab. 4: *RI* values

	n									
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
RI	0.00	0.00	0.58	0.90	1.12	1.24	1.32	1.41	1.45	1.49

Source: Jonuzi et al. (2024)

Step 2:

Composite weighted *Z*-score (CWZ_k) calculation (see Equation (7)) aggregating the contributions

Tab. 5: Used data sources to evaluate each sub criterion

Sub criterion *i.j*	Used data sources to evaluate sub criterion
1.1	Provincial and regional development indicators including economic and prosperity metrics
1.2	Data on small, medium, and large private trucking operators registered in candidate locations
1.3	Port area storage efficiency data, including warehouse storage and destination delivery performance
2.1	Thai and foreign workforce statistics in candidate locations
2.2	Highway networks, railway networks, and airport infrastructure capable of supporting regional traffic, assessed by network density relative to total provincial area
2.3	Development trends for nonpolluting and environmentally sustainable infrastructure, including renewable energy use, wastewater management, and waste reduction in ports
3.1	Strategic plans or projects focused on logistics development in candidate locations
3.2	Budget allocation for investment and annual budget expenditure for provinces and regions
4.1	Strategic plans or projects focused on future railway system development
4.2	Candidate location capability level for future railway system development
4.3	Candidate location capability level for existing equipment utilization
4.4	Candidate location capability level for connecting to existing railway networks and future railway network development plans
4.5	Candidate location capability level for utilizing public utility resources in candidate and adjacent areas
4.6	Level of various environmental impacts
5.1	Feasibility of land acquisition
5.2	Alignment with national and railway development plans
5.3	Potential for investment cost recovery
5.4	Potential for creating added value for the region through selected location utilization

Source: own

of all criteria, weighted according to their AHP derived importance:

$$CWZ_k = \sum_{j=1}^{m} w_j \cdot Z_{k,j} \tag{7}$$

where: $w_j = w_i \cdot w_{i.j}$ represents AHP weight of criterion j and m is the total number of criteria considered.

For promising yard locations, composite weighted Z-score calculations were further complemented by sensitivity analysis for the main criteria weights involving adjustments of these weights to ±10% and ±20% tolerance. Larger variations of weights were excluded to preserve the practical relevance of the analysis and ensure that the weight shifts remained within reasonable expert-judgment limits (Baydaş et al., 2024).

The GIS analysis incorporated an examination of infrastructure and economic activity for promising yard locations indicated by the AHP including railway, road, industrial facility, and warehouse density evaluation within a 50-kilometer radius around promising container yard locations. The data and satellite imagery were collected from OpenStreetMap as well as from Thai National Spatial Data Infrastructure provided by the Department of Highways, the Department of Rail Transport, and the National Statistical Office of Thailand, then processed using ArcGIS software, developed by Esri, and an open-source mapping program, QGIS, to obtain road and railway densities. Using industrial company registration data retrieved from the Ministry of Industry, warehouse and industrial clustering were identified with help of Kernel Density Estimation. The data collection and processing steps were carried out between October 2023 and May 2024. Furthermore, an overlay analysis for site suitability assessment was performed, which considered average slope and land use patterns to avoid the yard to be potentially located, for example in densely populated areas or agricultural land. Based on Equation (8), the average slope was derived through spatial modeling analysis utilizing digital elevation model (DEM):

$$Average\ slope = arctan\left(\frac{\Delta h}{\Delta d}\right) \tag{8}$$

where: Δ*h* represents the elevation difference and Δ*d* denotes the horizontal distance.

The DEM data were obtained from the Thai Geo-Informatics and Space Technology Development Agency and the Shuttle Radar Topography Mission operated by NASA, which provide high spatial resolution. Similarly, land use data were classified using Sentinel-2 and Landsat-8 satellite imagery, applying supervised classification techniques, such as random forest and maximum likelihood classification, to categorize land into industrial zones, agricultural lands, and community areas.

Based on Equation (9), the industrial area was calculated as:

$$Industrial\ area = \frac{Total\ industrial\ area}{Studied\ area} \cdot 100\% \tag{9}$$

Based on Equation (10), the agricultural area was computed using land classification maps, specifically for regions identified as agricultural zones:

$$Agricultural\ area = \frac{Total\ agricultural\ area}{Studied\ area} \cdot 100\% \tag{10}$$

Based on Equation (11), the community area was calculated as:

$$Community\ area = \frac{Total\ community\ area}{Studied\ area} \cdot 100\% \tag{11}$$

As in case of economic activity assessment of promising locations, the overlay analysis for suitability assessment was conducted in ArcGIS and QGIS software within the same time frame. Finally, based on the methodology described by Chang et al. (2017), a hot spot analysis of investment value and average freight volume, by employing Getis-Ord Gi* statistics, was carried out, revealing significant clustering of logistics operations within a 50-kilometer radius of promising yard locations.

3 Results and discussion

In Tab. 6, the main and sub criteria weights coming from AHP methodology, steps 1 and 2 are summarized.

It can be seen in Tab. 6 that among main criteria, investment economics emerged as the most heavily weighted criterion (32.5%) indicating that stakeholders view return on investment, long term financial sustainability, and regional value creation as critical determinants of infrastructure development. This aligns with conclusions drawn by Kaliszewski et al. (2020) and Correia and Melo (2021), who assert that profitability often trumps even technical feasibility in logistics planning. Within investment economics, the regional value creation potential (36.4%) is perceived as most impactful that, for example taken together with a preference for sustainable infrastructure development (54.5%) under the infrastructure criterion, reflects a shift in thinking of decision makers who are clearly looking beyond short-term cost benefit analyses and more towards longer term systemic benefits. This concurs with previous findings summarized for example in Khan et al. (2022) or Göçer et al. (2022). Investment economics is followed by the government support criterion (22.5%) indicating that government support and policy alignment play crucial roles in location optimization, confirming the findings of Li et al. (2023)

Tab. 6: Main and sub criteria weights

Main criterion i	w_i (%)	Sub criterion $i.j$	$w_{i,j}$ (%)
1	13.5	1.1	33.1
		1.2	25.3
		1.3	41.6
2	14.6	2.1	17.3
		2.2	28.2
		2.3	54.5
3	22.5	3.1	32.3
		3.2	67.7
4	16.9	4.1	7.4
		4.2	14.0
		4.3	7.8
		4.4	17.6
		4.5	36.9
		4.6	16.3
5	32.5	5.1	20.5
		5.2	21.7
		5.3	21.4
		5.4	36.4

Source: own

on the importance of public sector involvement in logistics infrastructure development.

Rather than relying solely on algorithmic weighting, responses from 20 experts across government agencies, industry associations, and logistics firms were incorporated within the AHP methodology. This diversity in viewpoints helps balance the theoretical with the practical. However, it also requires evaluation of judgment consistency using consistency ratios. These ratios are summarized in Tab. 7.

As all consistency ratios in Tab. 7 are ≤ 0.1 this indicates acceptable consistency in the judgments further supporting the reliability of the findings and confirming that judgments are internally coherent.

To translate qualitative assessments into actionable rankings, AHP weights from Tab. 6 were integrated with *Z*-score normalization across 18 railway stations returning rankings based on the composite weighted *Z*-score. The *Z*-score normalized values for each sub criterion applied by the secondary data analysis that encompass several key national policies and strategic plans are summarized in Tab. 8.

Then, the composite weighted *Z*-scores for all assessed stations are shown in Tab. 9, in descending order.

The results in Tab. 9 predicts Nata Station in Nong Khai province, with a composite weighted *Z*-score equal to 0.8775, to significantly outperform all other candidates. Stations at Bua Yai Junction and Kut Chik Junction, both located in Nakhon Ratchasima province, follow with identical composite weighted *Z*-scores of 0.5203. All top 3 rankings belong to the major railway corridor intersecting the surveyed area from south to north. While stations at Bua Yai Junction and Kut Chik Junction are located more inland of the studied area, the proximity of Nata Station to the Thai-Laos border, at 8.5 km, positions it advantageously for international

Tab. 7: Consistency of judgments based on consistency ratios

Person	*i*	1.*j*	2.*j*	3.*j*	4.*j*	5.*j*
1	0.050	0.053	0.057	0.051	0.051	0.046
2	0.060	0.065	0.062	0.046	0.049	0.061
3	0.050	0.060	0.050	0.048	0.048	0.066
4	0.050	0.069	0.049	0.059	0.046	0.045
5	0.060	0.069	0.057	0.068	0.063	0.055
6	0.050	0.064	0.045	0.050	0.051	0.059
7	0.050	0.065	0.055	0.059	0.048	0.046
8	0.060	0.063	0.048	0.069	0.055	0.062
9	0.050	0.045	0.054	0.068	0.058	0.048
10	0.060	0.065	0.056	0.065	0.058	0.068
11	0.060	0.063	0.046	0.067	0.048	0.051
12	0.050	0.053	0.045	0.054	0.060	0.049
13	0.050	0.046	0.061	0.049	0.056	0.048
14	0.060	0.063	0.064	0.066	0.046	0.046
15	0.050	0.045	0.062	0.045	0.069	0.047
16	0.060	0.065	0.050	0.063	0.051	0.068
17	0.060	0.047	0.053	0.061	0.061	0.055
18	0.060	0.063	0.053	0.054	0.053	0.058
19	0.050	0.061	0.066	0.048	0.048	0.045
20	0.060	0.062	0.048	0.068	0.067	0.057

Source: own

Tab. 8: Z-score values ($Z_{k,j}$) – Part 1

k^i_{ij}	1			2		3			4						5			
	1.1	**1.2**	**1.3**	**2.1**	**2.2**	**3.1**	**3.2**	**3.3**	**4.1**	**4.2**	**4.3**	**4.4**	**4.5**	**4.6**	**5.1**	**5.2**	**5.3**	**5.4**
1	−0.7	0.1	−0.9	1.4	−0.5	−0.2	0.9	−0.8	−0.7	−0.4	−0.5	−0.5	−0.5	−0.8	0.7	−0.2	−0.7	1.0
2	−0.2	0.5	−0.6	1.4	−0.5	0.2	0.9	−1.1	−0.5	−0.4	0.0	0.0	−0.5	−0.2	0.7	0.1	−0.6	0.1
3	−0.2	0.5	−0.6	1.4	−0.5	0.2	−0.1	−1.1	−0.5	−0.4	0.0	0.0	−0.5	−0.2	0.7	0.1	−0.6	0.1
4	−0.9	−0.8	−1.4	1.4	−0.9	−1.2	0.9	0.8	−1.0	−0.4	−1.0	−1.0	−0.6	−0.8	0.7	−1.1	−0.9	0.9
5	0.3	0.7	−0.3	−0.2	−0.1	0.6	0.4	−2.1	−0.1	0.1	0.3	0.3	−0.4	0.3	−1.0	0.9	0.6	0.8
6	−0.2	1.1	−0.4	−0.2	0.0	0.5	0.4	0.7	−0.4	−0.4	0.2	0.2	−0.5	0.1	−1.0	1.3	−0.1	−0.8
7	−1.3	0.6	−0.4	−0.2	−0.2	0.2	0.9	−0.9	−0.5	−0.4	0.0	0.0	−0.5	−0.1	−1.0	0.8	−0.5	−1.4
8	−0.8	0.4	−0.7	−0.2	−0.9	0.1	0.9	0.8	−0.5	−0.4	−0.2	−0.2	−0.5	−0.2	−1.0	−1.0	−0.5	1.2
9	−1.4	0.9	0.3	−0.2	−0.2	1.4	0.9	0.1	0.0	−0.1	1.1	1.1	−0.4	0.5	−1.0	0.6	0.4	−0.4
10	−0.1	−0.8	−1.6	−1.7	−1.4	−1.3	0.4	1.3	−1.1	−0.4	−1.3	−1.3	−0.6	−0.9	−2.7	−2.3	−0.6	1.1

Tab. 8: Z-score values ($Z_{k,j}$) – Part 2

k_{ij}^{i}	1			2		3			4						5			
	1.1	1.2	1.3	2.1	2.2	3.1	3.2	3.3	4.1	4.2	4.3	4.4	4.5	4.6	5.1	5.2	5.3	5.4
11	0.4	−0.2	−1.6	−0.2	−0.1	−0.3	1.1	−2.1	−0.6	0.0	0.3	0.3	−0.4	0.3	−0.6	−0.1	0.6	0.8
12	−1.1	1.8	1.3	−0.2	0.3	1.5	1.1	0.7	0.4	−0.5	2.4	2.4	0.4	2.8	−0.6	0.4	2.8	−1.1
13	−1.1	1.8	1.3	−0.2	0.3	1.5	1.1	0.7	0.4	−0.5	2.4	2.4	0.4	2.8	−0.6	0.4	2.8	−1.1
14	−1.1	1.8	1.3	−1.7	0.3	1.5	−2.1	0.7	0.4	−0.5	2.4	2.4	0.4	2.8	−2.6	0.4	2.8	−1.1
15	−1.1	1.8	1.3	−1.7	0.3	1.5	0.4	0.7	0.4	−0.5	2.4	2.4	0.4	2.8	−2.6	0.4	2.8	−1.1
16	0.5	−0.9	1.0	1.4	3.6	−1.4	1.1	0.6	−1.0	−0.5	−1.0	−1.0	−0.6	−0.6	1.4	4.2	0.1	−1.4
17	−1.1	1.8	1.3	−0.2	0.3	1.5	−0.8	0.7	0.4	−0.5	2.4	2.4	0.4	2.8	−0.6	0.4	2.8	−1.1
18	0.4	−0.2	−0.3	−0.2	−0.1	−0.3	−1.5	−2.1	−0.6	0.0	0.3	0.3	−0.4	0.3	−0.6	−0.1	0.6	0.8

Source: own

Tab. 9: Station rankings based on the composite weighted Z-score (CWZ_k)

Rank	Station	Composite weighted Z-score
1	Nata	0.8775
2	Bua Yai Junction	0.5203
3	Kut Chik Junction	0.5203
4	Ban Kradone	0.4423
5	Ban Koh	0.2503
6	Chana Chira Junction	0.1462
7	Bung Wai	0.0042
8	Non Sa-ard	−0.0107
9	Ban Takho	−0.0231
10	Roi Et	−0.0660
11	Maha Sarakham	−0.0819
12	Phon Thong	−0.1080
13	Nong Takai	−0.1569
14	Udon Thani	−0.1893
15	Nong Waeng	−0.1989
16	Saphan Mittraphap 2	−0.2141
17	Buri Ruea	−0.3885
18	Amnat Charoen	−0.7565

Source: own

trade confirming the findings of Wei and Dong (2019) on cross border logistics optimization. Top 3 rankings representing the most promising yard locations were further investigated using sensitivity analysis (see Figs. 2–4), GIS analysis incorporating the evaluation of infrastructure and economic activity (see Tab. 10), overlay analysis (see Tab. 11) and hot spot analysis of logistics operations clustering (see Fig. 5).

Fig. 2 depicts the sensitivity analysis for Nata Station, which achieved the highest ranking in the basic scenario with a composite weighted *Z*-score of 0.8775. Upon increasing the weight of the Investment Economics criterion, the predominant element at 32.5%, by 10%, the *Z*-score ascended to 0.9312. An additional 20% augmentation in the weight of this criterion raised the *Z*-score to 0.9760, solidifying the leading position of Nata Station. In contrast, a 10% reduction in the investment economics weight resulted in a modest decrease of the *Z*-score to 0.8213, while a 20% reduction led to a decline to 0.7658. Notwithstanding these reductions, Nata Station retained its position as the premier destination, illustrating its considerable resilience and strategic significance across diverse weight scenarios. For additional criteria, including Government Support and Railway Network, modifications in weights produced minimal effects, with *Z*-score swings within a ±0.03 range. This indicates that although these elements enhance overall suitability, the primary determinants for this site are investment possibility and regional value creation. The results correspond to the findings of Kaliszewski et al. (2020) and Correia and Melo (2021), who emphasize the precedence of economic viability in infrastructure siting decisions.

Fig. 3 shows the sensitivity analysis for Bua Yai Junction Station.

Bua Yai Junction Station, which attained a second-place ranking with a base *Z*-score of 0.5203, exhibited substantial variability according to sensitivity analysis. A 10% augmentation in the Railway Network criterion (16.9% baseline weight) elevated the *Z*-score to 0.5586, whilst a 20% augmentation boosted it to 0.5922. A 10% reduction in this criterion resulted in a *Z*-score of 0.4867, while a 20% decline yielded a *Z*-score of 0.4531. The alterations suggest that the rating of Bua Yai Junction Station is influenced by railway connectivity and accessibility factors, thereby corroborating previous studies by Fahim et al. (2022) and Ulutaş et al. (2020), which highlight the strategic

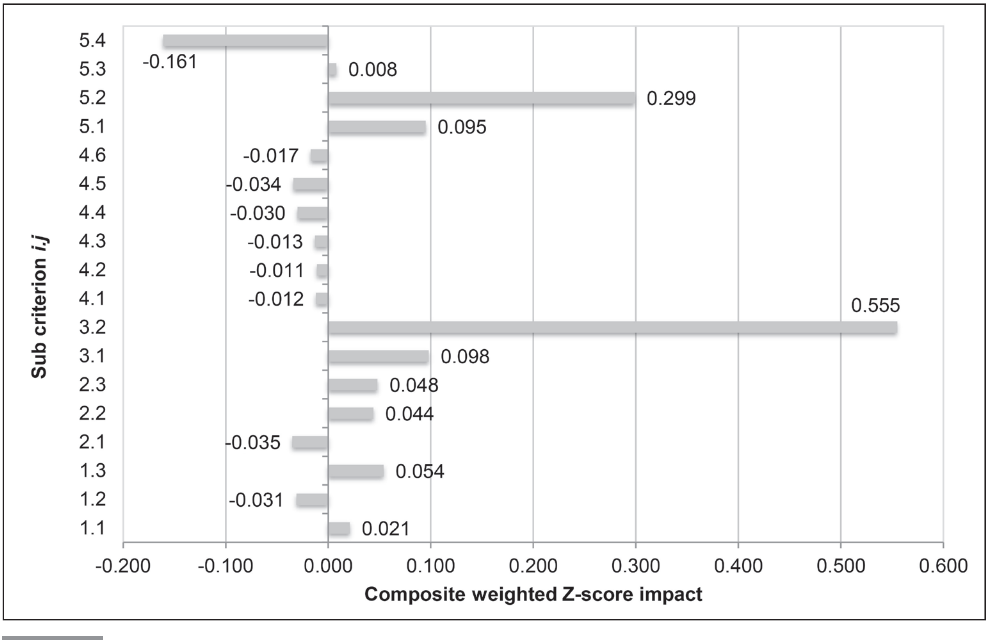

Fig. 2: Sensitivity analysis outputs for Nata Station

Source: own

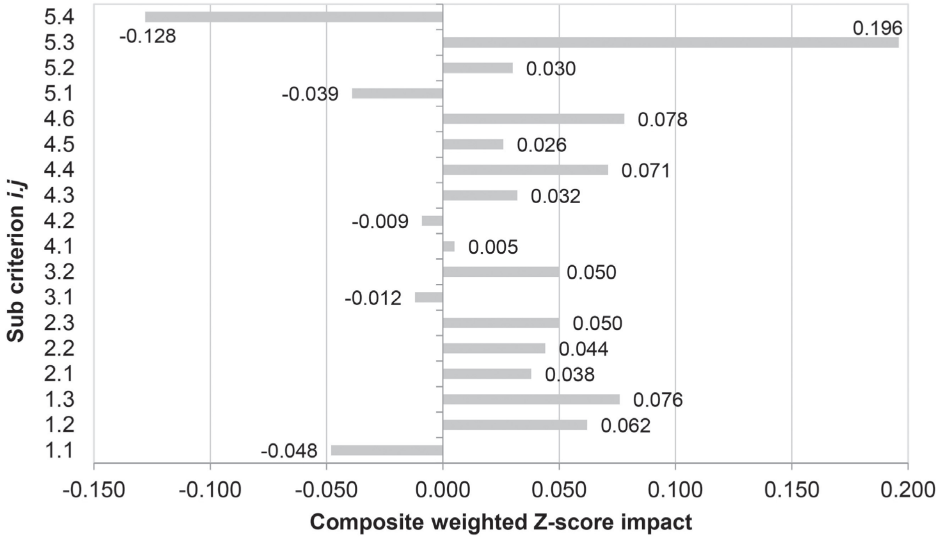

Fig. 3: Sensitivity analysis outputs for Bua Yai Junction Station

Source: own

significance of railway infrastructure in multimodal transport integration. Nevertheless, the augmentation of the Investment Economics weight resulted in just a marginal enhancement of the *Z*-score for the station from 0.5203 to 0.5427, indicating that the its fundamental advantage is in transport connection rather than investment-driven potential.

Finally, Fig. 4 summarizes the outputs of the sensitivity analysis for Kut Chik Junction Station.

Kut Chik Junction Station, sharing second place with Bua Yai Junction Station (base *Z*-score 0.5203), exhibited analogous tendencies in response to weight change. A 10% increase in the Government Support criterion elevated its *Z*-score to 0.5539, while a 20% increase resulted in a *Z*-score of 0.5865. Reducing the same criterion by 10% or 20% resulted in *Z*-score drops to 0.4916 and 0.4628, respectively. Notably, modifications to the infrastructure weight had negligible effects (less than ±0.02), suggesting that although infrastructure holds significance, it is not the main differentiator for Kut Chik Junction Station. This underscores that the alignment of public policy and financial backing are essential for realizing the complete potential of the site, as shown by Liu et al. (2021) and Göçer et al. (2022).

Tab. 10 shows the infrastructure and economic activity in promising locations as densities of the following resources: road, railway, industrial facility, and warehouse.

In all resources stated in Tab. 10, Nata Station outperforms the stations located at Bua Yai Junction and Kut Chik Junction. It has the highest road and railway densities of 0.85 km/km² and 0.32 km/km², respectively enhancing multimodal transport integration and accessibility, which are key advantages in container yard operations as other scholars have concluded, for example in Moskvichev et al. (2021). Furthermore, as summarized in Tab. 11, Nata Station also has the best physical suitability with the greatest industrial land share (i.e., 68%) and the lowest agricultural use (i.e., 8%).

Combined with a gentle average slope of just 2.3%, the area is both amenable to construction and minimally disruptive. These are exactly the kind of terrain and land use dynamics emphasized by Herzog (2021), who underscores the importance of practical geography in the deployment of infrastructure. Spatial clustering analysis using Getis-Ord Gi*

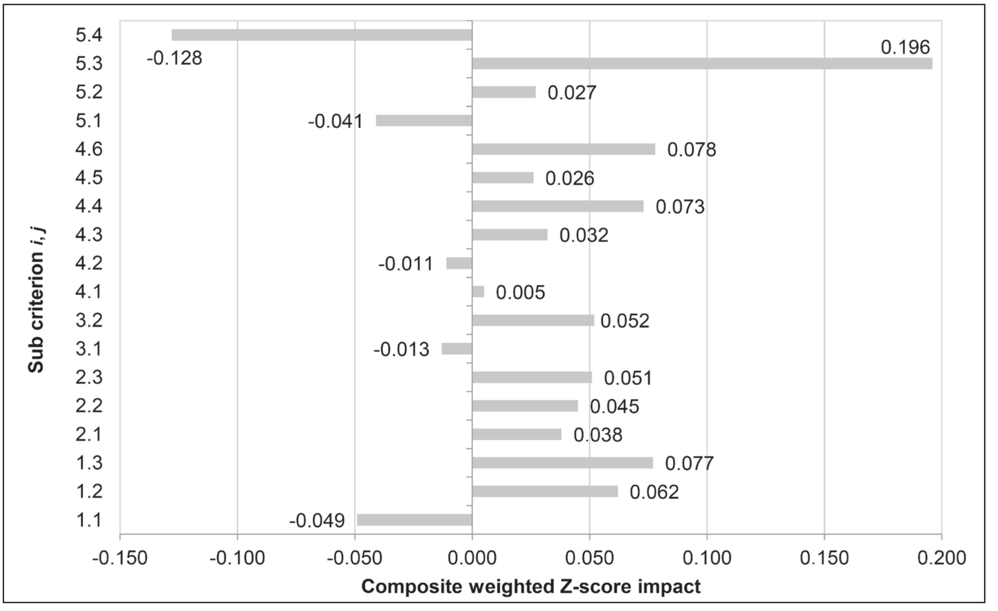

Fig. 4: **Sensitivity analysis outputs for Kut Chik Junction Station**

Source: own

Tab. 10: **Infrastructure and economic activity in promising locations**

Station	Density			
	Road (km/km^2)	Railway (km/km^2)	Industrial facility (facilities/km^2)	Warehouse (facilities/km^2)
Nata	0.85	0.32	2.8	1.5
Bua Yai Junction	0.78	0.28	2.3	1.2
Kut Chic Junction	0.72	0.25	2.1	1.1

Source: own

Tab. 11: **Overlay analysis of promising locations**

Factor	Station		
	Nata	Bua Yai Junction	Kut Chik Junction
Average slope (%)	2.3	3.1	3.5
Industrial area (%)	68	55	52
Community area (%)	12	15	18
Agricultural area (%)	8	12	15

Source: own

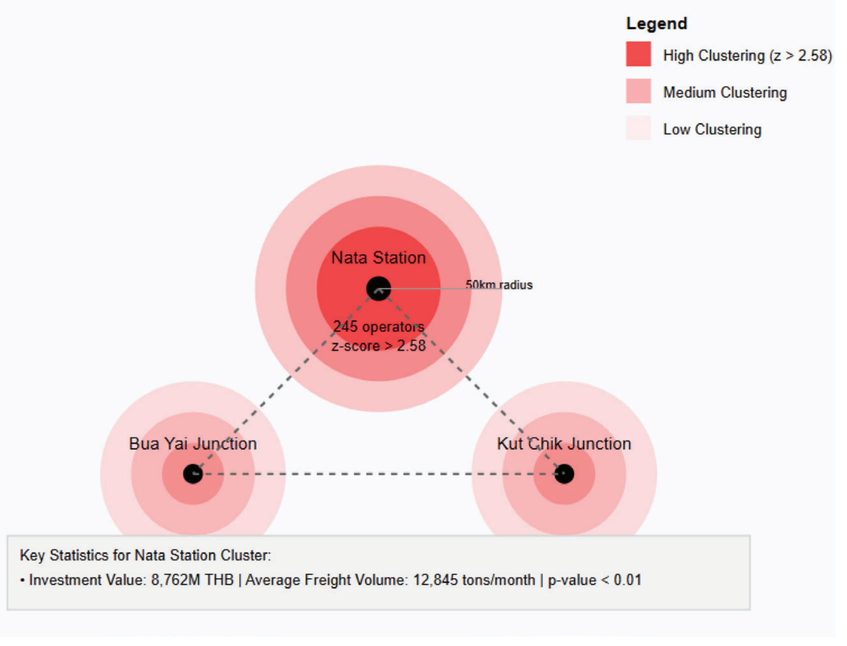

Fig. 5: Hot spot analysis of logistics operations clustering around promising locations

Source: own

statistics (see Fig. 5) offers a final layer of insight.

The concentration of 245 logistics operators within the 50km radius around Nata Station, with total investments of 8,762 million Thai baht demonstrates not only its strategic potential but also its status, in terms of functional characteristics, as an existing logistics hub, surrounded by dense industrial and warehousing activity. This pre-existing ecosystem strengthens the business case for investment, mirroring the arguments of De Langen et al. (2020), who emphasize the pull effect of logistics clusters in attracting capital and ancillary industries.

Conclusions

This study examined the optimal location for container yard development in Northeastern Thailand to enhance multimodal transport connectivity and strengthen national competitiveness in regional and global trade. Utilizing an AHP and GIS based framework, the study evaluated 18 railway stations based on five key factors: logistics suitability (13.5%), infrastructure readiness (14.6%), government support (22.5%), railway network capabilities (16.9%), and investment economics (32.5%). The findings identified Nata Station in Nong Khai Province as the most strategically advantageous location, with the highest composite weighted *Z*-score of 0.8775. The station features a road network density of 0.85 km per square kilometer and a railway density of 0.32 km per square kilometer, facilitating efficient cross border freight transportation. The establishment of a container yard in this optimal location is expected to reduce logistics costs, improve supply chain efficiency, and enhance cross border trade, particularly with member states of the Association of South East Asian Nations (ASEAN) trading bloc and China. This development reinforces the role of Thailand as a regional logistics hub, fostering industrial expansion and economic integration. Among all sub-criteria, three emerged as the most influential drivers of the final ranking: (1) regional value creation potential (5.4, composite weight 11.83%), reflecting the strategic importance of long-term economic benefits for the surrounding region; (2) investment return potential (5.3, composite weight 6.96%), highlighting the primacy of financial viability in infrastructure decision-making; and (3) sustainable infrastructure development (2.3, composite weight 7.96%), underscoring the growing emphasis on environmental responsibility in logistics planning. These findings are particularly significant in the context of developing regional logistics resources in Northeastern Thailand. The top-ranked location, Nata Station in Nong Khai Province, benefits directly from its position on the Trans-Asian Railway corridor and proximity to the Laos-China Railway link, reinforcing its role as a gateway to regional economic corridors including the Greater Mekong Subregion Economic Cooperation Program and the Pan-Asian Railway Network. The strategic concentration of investment potential at this location directly supports the national logistics competitiveness of Thailand by reducing supply-chain costs and improving multimodal transport connectivity with ASEAN and Chinese markets. In addition, based on analysis by the Bank of Thailand (2025), Nong Khai province has exhibited relatively modest and fluctuating economic growth compared to the national average, with GRP growth reported at approximately 0.7%–1.7% in 2024 and 1.2%–2.2% in 2025. This trend reflects the structural characteristics of the region, including a strong dependence on agriculture, relatively low industrial concentration, and weaker household purchasing power, which contribute to ongoing regional disparities and labor migration trends. Therefore, infrastructure investment at Nata Station can support regional development by attracting logistics activities, creating jobs, and reducing migration pressures. Government support plays an important role in enabling such development through policy alignment and investment incentives.

The findings of this study carry direct implications for logistics infrastructure policy in Thailand. The identification of Nata Station as the optimal container yard location provides an evidence-based recommendation for the State Railway of Thailand, the Office of Transport and Traffic Policy and Planning, and private investors considering multimodal freight infrastructure in the northeastern region. The AHP-GIS framework developed here can be readily adapted for similar site-selection tasks in other Thai regions or comparable developing economies. Policymakers should prioritize investment incentives, land-use zoning, and railway capacity upgrades at and around Nata Station to accelerate the realization of its regional value creation potential. The study also underscores that government support and cross-border connectivity are critical enablers of national logistics competitiveness, reinforcing the case for continued public-private partnership in multimodal transport corridor development.

While the AHP and GIS based framework proposed in this work provides a strategic tool for policymakers and private sector stakeholders to leverage logistics infrastructure and strengthen long term competitiveness in global trade, it has several limitations that should be acknowledged. First, AHP relies on expert judgment, which is inherently subjective; although consistency ratios confirmed the reliability of responses, the results may vary with different expert panels. Second, the *Z*-score normalization approach assumes linear relationships between criteria scores and overall suitability, which may not fully capture non-linear trade-offs. Third, the secondary data used for sub-criterion evaluation reflects conditions at the time of data collection (2023–2024) and may not account for future infrastructure changes or policy shifts. Fourth, the GIS overlay analysis is limited by the resolution of available satellite imagery and administrative data.

Future development of the proposed strategic tool should consider integrating dynamic decision-making frameworks that account for time-sensitive variables such as land value fluctuations, evolving transport technology, and regulatory change. The analytic network process could be applied where interdependencies among criteria are complex. Fuzzy logic extensions to AHP may improve robustness when dealing with qualitative expert data. Additionally, simulation-based optimization and multi-objective modeling could complement the present framework by optimizing yard layout, storage configuration, and operational resource allocation at the selected site.

References

Abouhawwash, M., & Jameel, M. (2023). Evaluation factors of solar power plants to reduce cost under neutrosophic multi-criteria decision making model. *Sustainable Machine Intelligence Journal*, *2*(1), 1–11. https://doi.org/10.61185/SMIJ.2023.22101

Abu Aisha, T., Ouhimmou, M., & Paquet, M. (2020). Optimization of container terminal layouts in the seaport – Case of port of Montreal. *Sustainability*, *12*(3), 1165. https://doi.org/10.3390/su12031165

Aldrighetti, R., Battini, D., Ivanov, D., & Zennaro, I. (2021). Costs of resilience and disruptions in supply chain network design models: A review and future research directions. *International Journal of Production Economics*, *235*, 108103. https://doi.org/10.1016/j.ijpe.2021.108103

Bandyopadhyay, A., & Bhatnagar, S. (2023). Impact of COVID-19 on ports, multimodal logistics and transport sector in India: Responses and policy imperatives. *Transport Policy*, *130*, 15–25. https://doi.org/10.1016/j.tranpol.2022.10.014

Bank of Thailand. (2025). *Regional economic outlook: Gross regional product (GRP) of Northeastern Thailand (Isan region)*. https://www.bot.or.th/en/thai-economy/economic-outlook.html

Baydaş, M., Elma, O. E., & Stević, Ž. (2024). Proposal of an innovative MCDA evaluation methodology: Knowledge discovery through rank reversal, standard deviation, and relationship with stock return. *Financial Innovation*, *10*(1), 1–35. https://doi.org/10.1186/s40854-023-00526-x

Belay, S., Goedert, J., Woldesenbet, A., & Rokooei, S. (2022). AHP based multi criteria decision analysis of success factors to enhance decision making in infrastructure construction projects. *Cogent Engineering*, *9*(1), 1–15. https://doi.org/10.1080/23311916.2022.2043996

Bhuvaneskumar, A., Sivakumar, V. J., & Pushparaj, N. (2023). Performance assessment and ranking of socially responsible companies in India using FAHP, TOPSIS and Altman *Z*-score. *Benchmarking: An International Journal*, *30*(3), 736–765. https://doi.org/10.1108/BIJ-09-2021-0512

Bui, V. D., & Nguyen, H. P. (2022). The role of the inland container depot system in developing a sustainable transport system. *International Journal of Knowledge-Based Development*, *12*(3/4), 424–443. https://doi.org/10.1504/ijkbd.2022.128914

Cahyono, Y., Purwoko, D., Koho, I. R., Setiani, A., Supendi, S., Setyoko, P. I., Sosiady, M., & Wijoyo, H. (2023). The role of supply chain management practices on competitive advantage and performance of halal agroindustry SMEs. *Uncertain Supply Chain Management*, *11*(1), 153–160. https://doi.org/10.5267/j.uscm.2022.10.012

Chang, B. A., Pearson, W. S., & Owusu-Edusei, K. (2017). Correlates of county-level nonviral sexually transmitted infection hot spots in the US: Application of hot spot analysis and spatial logistic regression. *Annals of Epidemiology*, *27*(4), 231–237. https://doi.org/10.1016/j.annepidem.2017.02.004

Correia, I., & Melo, T. (2021). Integrated facility location and capacity planning under uncertainty. *Computational and Applied Mathematics*, *40*(5), 1–36. https://doi.org/10.1007/s40314-021-01560-0

Curry, G. N., Nake, S., Koczberski, G., Oswald, M., Rafflegeau, S., Lummani, J., Peter, E., & Nailina, R. (2021). Disruptive innovation in agriculture: Socio-cultural factors in technology adoption in the developing world. *Journal of Rural Studies*, *88*, 422–431. https://doi.org/10.1016/j.jrurstud.2021.07.022

Dai, Y. (2022). China's infrastructure investment to the belt and road: The case of the China-Indochina Peninsula economic corridor. *The Chinese Economy*, *55*(3), 169–187. https://doi.org/10.1080/10971475.2021.1958452

Danyluk, M. (2019). Fungible space: Competition and volatility in the global logistics network. *International Journal of Urban and Regional Research*, *43*(1), 94–111. https://doi.org/10.1111/1468-2427.12675

De Langen, P. W., Sornn-Friese, H., & Hallworth, J. (2020). The role of port development companies in transitioning the port business ecosystem; the case of port of Amsterdam's circular activities. *Sustainability*, *12*(11), 4397. https://doi.org/10.3390/su12114397

Fahim, P. B. M., Rezaei, J., Montreuil, B., & Tavasszy, L. (2022). Port performance evaluation and selection in the physical internet. *Transport Policy*, *124*, 83–94. https://doi.org/10.1016/j.tranpol.2021.07.013

Garrow, L. A., German, B. J., & Leonard, C. E. (2021). Urban air mobility: A comprehensive review and comparative analysis with autonomous and electric ground transportation for informing future research. *Transportation Research Part C: Emerging Technologies*, *132*, 103377. https://doi.org/10.1016/j.trc.2021.103377

Göçer, A., Özpeynirci, Ö., & Semiz, M. (2022). Logistics performance index-driven policy development: An application to Turkey. *Transport Policy*, *124*, 20–32. https://doi.org/10.1016/j.tranpol.2021.03.007

Herzog, I. (2021). National transportation networks, market access, and regional economic growth. *Journal of Urban Economics*, *122*, 103316. https://doi.org/10.1016/j.jue.2020.103316

Jiang, C., Wan, Y., Yang, H., & Zhang, A. (2021). Impacts of high-speed rail projects on CO^2 emissions due to modal interactions: A review. *Transportation Research Part D: Transport and Environment*, *100*, 103081. https://doi.org/10.1016/j.trd.2021.103081

Jiao, J., Wang, J., Zhang, F., Jin, F., & Liu, W. (2020). Roles of accessibility, connectivity and spatial interdependence in realizing the economic impact of high-speed rail: Evidence from China. *Transport Policy*, *91*, 1–15. https://doi.org/10.1016/j.tranpol.2020.03.001

Jonuzi, E., Alkan, T., Durduran, S. S., & Selvi, H. Z. (2024). Using GIS-supported MCDA method for appropriate site selection of parking lots: The case study of the city of Tetovo, North Macedonia. *International Journal of Engineering and Geosciences*, *9*(1), 86–98. https://doi.org/10.26833/ijeg.1319605

Kaliszewski, A., Kozłowski, A., Dąbrowski, J., & Klimek, H. (2020). Key factors of container port competitiveness: A global shipping lines perspective. *Marine Policy*, *117*, 103896. https://doi.org/10.1016/j.marpol.2020.103896

Kayikci, Y. (2021). Stream processing data decision model for higher environmental performance and resilience in sustainable logistics infrastructure. *Journal of Enterprise Information Management*, *34*(1), 140–167. https://doi.org/10.1108/jeim-08-2019-0232

Khan, S. A. R., Godil, D. I., Yu, Z., Abbas, F., & Shamim, M. A. (2022). Adoption of renewable energy sources, low-carbon initiatives, and advanced logistical infrastructure – An step toward integrated global progress. *Sustainable Development*, *30*(1), 275–288. https://doi.org/10.1002/sd.2243

Khaslavskaya, A., & Roso, V. (2020). Dry ports: Research outcomes, trends, and future implications. *Maritime Economics & Logistics*, *22*(2), 265–292. https://doi.org/10.1057/s41278-020-00152-9

Khorram, S. (2020). A novel approach for ports' container terminals' risk management based on formal safety assessment: FAHP-entropy measure – VIKOR model. *Natural Hazards*, *103*(2), 1671–1707. https://doi.org/10.1007/s11069-020-03976-z

Kramarz, M., & Przybylska, E. (2021). Multimodal transport in the context of sustainable development of a city. *Sustainability*, *13*(4), 2239. https://doi.org/10.3390/su13042239

Li, L., Wang, J., Wang, H., Jin, X., & Du, L. (2023). Intermodal transportation hub location optimization with governments subsidies under the Belt and Road Initiative. *Ocean & Coastal Management*, *231*, 106414. https://doi.org/10.1016/j.ocecoaman.2022.106414

Li, S., Cao, X., Liao, W., & He, Y. (2020). Factors in the sea ports-of-entry and road ports-of-entry cross-border logistics route choice. *Journal of Transport Geography*, *84*, 102689. https://doi.org/10.1016/j.jtrangeo.2020.102689

Liu, J., Wang, X., & Guo, J. (2021). Port efficiency and its influencing factors in the context of pilot free trade zones. *Transport Policy*, *105*, 67–79. https://doi.org/10.1016/j.tranpol.2021.02.011

Lo Storto, C., & Evangelista, P. (2023). Infrastructure efficiency, logistics quality and environmental impact of land logistics systems in the EU: A DEA-based dynamic mapping. *Research in Transportation Business & Management*, *46*, 100814. https://doi.org/10.1016/j.rtbm.2022.100814

Moskvichev, O., Moskvicheva, E., & Bulatov, A. (2021). Clustering methods for determination of optimal locations of container storage and distribution centers. *Transportation Research Procedia*, *54*, 461–469. https://doi.org/10.1016/j.trpro.2021.02.096

Pazzini, M., Corticelli, R., Lantieri, C., & Mazzoli, C. (2022). Multi-criteria analysis and decision-making approach for the urban regeneration: The application to the Rimini Canal Port (Italy). *Sustainability*, *15*(1), 772. https://doi.org/10.3390/su15010772

Prasittisopin, L., Kitkuakul, P., Chotchakornpant, K., & Rugkhapan, N. T. (2024). Implementing the sister city policy: Perspectives from Thailand. *Nakhara: Journal of Environmental Design and Planning*, *23*(2), 413. https://doi.org/10.54028/NJ202423413

Puchongkawarin, C., & Ransikarbum, K. (2021). An integrative decision support system for improving tourism logistics and public transportation in Thailand. *Tourism Planning & Development*, *18*(6), 614–629. https://doi.org/10.1080/21568316.2020.1837229

Sakai, T., Beziat, A., & Heitz, A. (2020). Location factors for logistics facilities: Location choice modeling considering activity categories. *Journal of Transport Geography*, *85*, 102710. https://doi.org/10.1016/j.jtrangeo.2020.102710

Song, D. (2021). A literature review, container shipping supply chain: Planning problems and research opportunities. *Logistics*, *5*(2), 41. https://doi.org/10.3390/logistics5020041

Tadić, S., Krstić, M., Roso, V., & Brnjac, N. (2020). Dry port terminal location selection by applying the hybrid grey MCDM model. *Sustainability*, *12*(17), 1–24. https://doi.org/10.3390/su12176983

Tavana, M., Soltanifar, M., & Santos-Arteaga, F. J. (2023). Analytical hierarchy process: Revolution and evolution. *Annals of Operations Research*, *326*(2), 879–907. https://doi.org/10.1007/s10479-021-04432-2

Theerathitichaipa, K., Wisutwattanasak, P., Se, C., Seefong, M., Jomnonkwao, S., Champahom, T., Ratanavaraha, V., & Kasemsri, R. (2024). Assessment of disparity in accessing railway stations in Thailand: An application geographic information system network analysis. *Journal of Geovisualization and Spatial Analysis*, *8*(1), 1–18. https://doi.org/10.1007/s41651-023-00168-8

Ulutaş, A., Karakuş, C. B., & Topal, A. (2020). Location selection for logistics center with fuzzy SWARA and CoCoSo methods. *Journal of Intelligent & Fuzzy Systems*, *38*(4), 4693–4709. https://doi.org/10.3233/JIFS-191400

Wangai, A. W., Rohacs, D., & Boros, A. (2020). Supporting the sustainable development of railway transport in developing countries. *Sustainability*, *12*(9), 3572. https://doi.org/10.3390/su12093572

Wei, H., & Dong, M. (2019). Import-export freight organization and optimization in the dry-port-based cross-border logistics network under the Belt and Road Initiative. *Computers & Industrial Engineering*, *130*, 472–484. https://doi.org/10.1016/j.cie.2019.03.007

Xiong, W., Chen, B., Wang, H., & Zhu, D. (2020). Public-private partnerships as a governance response to sustainable urbanization: Lessons from China. *Habitat International*, *95*, 102095. https://doi.org/10.1016/j.habitatint.2019.102095

Xu, Y., Liu, S., Wang, J., & Shang, X. (2022). Consensus checking and improving methods for AHP with q-rung dual hesitant fuzzy preference relations. *Expert Systems with Applications*, *208*, 117902. https://doi.org/10.1016/j.eswa.2022.117902

Yannis, G., Kopsacheili, A., Dragomanovits, A., & Petraki, V. (2020). State-of-the-art review on multi-criteria decision-making in the transport sector. *Journal of Traffic and Transportation Engineering*, *7*(4), 413–431. https://doi.org/10.1016/j.jtte.2020.05.005

Zhou, X., Wei, X., Lin, J., Tian, X., Lev, B., & Wang, S. (2021). Supply chain management under carbon taxes: A review and bibliometric analysis. *Omega*, *98*, 102295. https://doi.org/10.1016/j.omega.2020.102295

Do investors really care about sustainability? Evidence from European companies on the relationship between ESG performance and stock liquidity

Sebnem Yasar[1], Hakan Ozkaya[2]

[1] *Dokuz Eylul University, Izmir Vocational School, Department of Accounting and Tax Applications, Turkey, ORCID: 0000-0001-6173-5148, sebnem.yasar@deu.edu.tr (corresponding author);*
[2] *Dokuz Eylul University, School of Applied Sciences, Department of International Trade, Turkey, ORCID: 0000-0001-7810-5943, hakan.ozkaya@deu.edu.tr.*

Abstract: *This study examines the impact of corporate environmental, social and governance (ESG) performance on capital market efficiency, namely stock liquidity. Drawing on a dataset of 1,693 publicly traded firms across 23 European stock exchanges between 2002 and 2022, we investigate whether companies with stronger ESG credentials benefit from more liquid equity markets. Our results provide strong evidence that ESG performance has a significant positive effect on stock liquidity. This suggests that a substantial portion of European investors today prefer to invest in companies that act environmentally friendly, value social rights and equality, and engage in good governance. Among the three ESG dimensions, the environmental component exhibits the most significant effect, indicating that investors particularly value firms that engage in environmentally responsible practices. Furthermore, the analysis reveals notable regional differences in the importance assigned to ESG factors, showing that investors in Western and Northern European countries show a stronger preference for sustainability-oriented firms compared to their counterparts in Eastern and Southern Europe. These findings highlight the influence of regional economic, cultural, and regulatory contexts on investment behavior. Overall, our study contributes to the growing literature on sustainable finance by underscoring the role of ESG performance not only as a tool for ethical or reputational enhancement but also as a mechanism that can directly improve financial market outcomes. By identifying ESG performance as a determinant of stock liquidity, this research also supports the integration of ESG considerations into investment decision-making and corporate strategy.*

Keywords: *Stock liquidity, ESG, sustainability, European stock markets, investor behavior.*

JEL Classification: *M14, G10, G11, Q01.*

APA Style Citation: Yasar, S., & Ozkaya, H. (2026). Do investors really care about sustainability? Evidence from European companies on the relationship between ESG performance and stock liquidity. *E&M Economics and Management, 29*(2), 177–196. https://doi.org/10.15240/tul/001/2026-2-012

Introduction

The asset's market liquidity indicates the ability of investors to buy and sell substantial amounts of assets quickly, at low cost, and without significant price concessions. Amihud et al. (2006) stated that stock market liquidity indicates the existence of willing buyers and sellers who agree to exchange a certain amount of securities at the specified price without any delay. As a fundamental market feature, its presence secures the smooth functioning of the market and its absence creates unrest in it. Stock liquidity is closely related to market recognition and largely reflects how corporate decisions and activities are evaluated by market actors. Therefore, companies' business missions, principles, and values are represented by stock market liquidity to a great extent, and environmental, social and governance (ESG) philosophy is no exception.

Global warming and climate change have revealed the importance of developing environmental protection awareness, and in this regard, the roles and responsibilities of companies have begun to be discussed. Today, corporate information on issues related to greenhouse gas emissions, energy and water consumption, recycling or disposal of waste materials, and actions of the companies to decrease the negative effects of their activities on nature are of foremost significance. In addition, companies are expected to attach utmost importance to issues related to human rights and gender equality, take part in social responsibility activities, and always prioritize transparency and accountability in corporate governance processes.

ESG is a non-financial business evaluation system focusing on environmental, social, and governance performance, and aims to measure the sustainable development ability of firms and determine whether they have adequate corporate environmental, social, and governance responsibility awareness. ESG was created by the United Nations Global Compact in 2004, and with this concept, businesses were required to give due attention to environmental protection, social responsibility, and corporate governance in their operations and business development processes and to quantify their non-financial performances. The purpose of conducting ESG assessment is to encourage businesses to invest in ESG and to help them integrate ESG priorities into their strategic decision-making processes and allocate their resources more effectively.

Recent studies increasingly emphasize that stock liquidity is influenced not only by financial indicators but also by non-financial performance factors such as ESG. ESG is increasingly recognized by all stakeholders, especially investors, and companies' ESG ratings are increasingly taken into account in decision-making processes (Bihari et al., 2025; Bhowmick et al., 2025; Hutchings, 2025; Sanseverino et al., 2024). Companies with strong ESG performance typically enjoy higher investor confidence, lower transaction costs, and better market accessibility (Broadstock et al., 2021; Chen & Xie, 2022; S. Chen et al., 2023; Mohammad et al., 2021). As a result, more and more firms are trying to disclose their ESG activities.

The purpose of this study is to examine the effect of ESG performance on capital market performance. The main research question is whether ESG performance improves stock liquidity in European capital markets. Specifically, we examine whether higher ESG performance leads to higher stock liquidity, whether this effect varies across European regions, and through which possible channels ESG performance may influence stock liquidity. For this purpose, we test the relationship between ESG scores and the stock liquidity of European companies. Although there are different views, the majority of the literature shows that ESG activities and disclosures decrease information asymmetry between the firms and the investors, enhance corporate reputation, decrease firm risk, and attract more attention and interest from institutional investors, all of which affect stock liquidity.

Our dataset includes European listed companies between 2002–2022 with ESG scores obtained from Refinitiv Eikon database. To ensure the robustness of our findings we conduct a series of robustness tests, including alternative model specifications and control variables.

Our study contributes to literature in several ways. First, there is no comprehensive study in the literature investigating the relationship between ESG performance and stock liquidity for European companies. Therefore, this study fills this gap and enriches the literature by shedding light on the impact of ESG scores on stock liquidity. Second, it extends studies on factors affecting stock liquidity by identifying ESG performance as an impact factor in capital markets. Third, this study provides empirical support for encouraging ESG investments by revealing the positive impact of ESG performance on performance in capital markets. Finally, this study reveals regional differences in investors' reactions to ESG scores, providing insight into the regions where efforts to increase environmental, social and/or governmental responsibility awareness should be concentrated.

The remaining parts of the paper are organized as follows. Section 1 includes a comprehensive literature review, and the hypothesis of the study developed in this context. Section 2 outlines the data of the research, and defines the research methodology. Section 3 presents the results of the analysis and the robustness tests, and discusses the findings.

1 Theoretical background and hypothesis development

1.1 ESG and stock liquidity

There are a few studies in the literature examining the impact of ESG disclosures on stock liquidity, and in all of these studies, a positive relationship between ESG performance and stock liquidity was found (Aboud & Diab, 2019; R. Chen et al., 2023; He et al., 2023; Liang et al., 2023; Luo, 2022; Meng-tao et al., 2023; Subramaniam et al., 2016; Wang et al., 2023; Zhang et al., 2024). Wang et al. (2023) stated that ESG performance increases firms' stock liquidity by reducing corporate risks and increasing stakeholders' support. Meng-tao et al. (2023) attributed the increase in stock liquidity to the decrease in corporate risks and the increase in institutional investors. R. Chen et al. (2023) found that good ESG performance increases stock liquidity by reducing agency costs, increasing foreign ownership, and improving corporate reputation. Zhang et al. (2024) suggested that high ESG scores lead to high stock liquidity by leading to an increase in corporate transparency and risk mitigation.

In addition to the studies mentioned above, literature examining the relationship between ESG, financial performance and firm value can also be used to understand the dynamics of the relationship between ESG performance and stock liquidity.

1.2 ESG, financial performance, and firm value

The question of how ESG ratings affect companies' financial performance and firm values is a controversial issue, often based on one of two contrasting theoretical perspectives: (i) the shareholder theory; and (ii) the stakeholder theory.

(i) Shareholder theory: the first of these is Friedman's shareholder theory, which argues that the company's only responsibility is to its shareholders. Friedman (1970) criticized the increasing attention to corporate social responsibility among companies and the academic community, stating that the only goal of a company is maximizing its profits, and any activity carried out under the name of social responsibility will create conflict between different stakeholders of the company. From this perspective, investments in ESG are regarded as irrational decisions that undermine shareholders' interests by creating more virtuous companies at higher costs, which in turn negatively affect financial performance, and are likely to be perceived unfavorably by existing and potential investors.

(ii) Stakeholder theory: the second theory is Freeman's stakeholder theory, which argues that the company's responsibility is not only to its shareholders but to all its stakeholders. Freeman (1984) defined stakeholders as any group or individual who can affect a company's achievement of its goal or is affected by the company's actions and stated that the true success of a company depends on satisfying all its stakeholders. According to this perspective, social responsibility has both economic and moral dimensions, and companies that can effectively manage these dimensions will gain competitive advantage. In this regard, it can be expected that capital markets will not remain unresponsive to the ESG scores of companies that increase their corporate reputation and reliability by considering the demands of their stakeholders and will reward high performance.

In parallel with these two views, which are in complete contradiction with each other, studies in literature also exhibit contradictory results. Much of the previous literature has suggested that ESG disclosures help improve financial performance (Aboud & Diab, 2019; Kao, 2023; Lins et al., 2017; Malik & Kashiramka, 2024; Margolis, 2009; Noja et al., 2024; Rossi & Harjoto, 2020; Ye et al., 2022; Yoo et al., 2022) and firm value (Aydoğmuş et al., 2022; Bagh et al., 2024; Fatemi et al., 2015; Krüger, 2015; Miralles-Quirós et al., 2018; Sassen et al., 2016). However, some studies have found that non-financial information disclosures reduce financial performance and firm value (Alareeni & Hamdan, 2020; Brammer et al., 2006; Chen et al., 2018; Duque-Grisales & Aguilera-Caracuel, 2021; Khan & Liu, 2023). Agarwala et al. (2024) have revealed the existence of a U-shaped relationship between ESG scores and firm value meaning that the firms must exceed a certain threshold for a positive relationship to occur. Fatemi et al. (2018) have stated that ESG activities increase firm value, but ESG disclosures reduce firm value. Saygili et al. (2022) found that while social and governance disclosures have positive effect, environmental disclosures have negative effect on financial performance. Liu et al. (2022) have found out that social dimension has a more significant effect on financial performance of energy companies compared to the other two dimensions. Some other studies have argued that there is no relationship between ESG, financial performance, and firm value (Atan et al., 2018; Horvathova, 2010; Nag & Bhattacharyya, 2016). Capelle-Blancard and Petit (2019) have found that investors react negatively to negative ESG performance, but the reaction to good performance is insignificant.

1.3 Mechanisms linking ESG and stock liquidity

High ESG performance is expected to alleviate information asymmetry and agency conflicts, reduce corporate risks and improve risk management through better corporate governance practices, enhance corporate reputation, and attract institutional investors. As a result, good ESG performance has the potential to increase stock liquidity in capital markets by positively affecting financial performance and firm value.

According to agency theory, the separation of ownership and control in companies creates an agency relationship, and the agency relationship creates information asymmetry between shareholders and managers. With more information, management has the opportunity to make decisions against investors, and this leads to conflicts of interest and agency costs between the principal and the agent (Jensen & Meckling, 1976). Diamond and Verrecchia (1991) revealed that information asymmetry creates implicit costs between companies and the market and can lead to misallocation of resources, and policies that reduce this asymmetry can increase stock liquidity. When investors buy or sell shares, they rely on the private information they have about the company. This information may be financial or non-financial information such as ESG disclosures. ESG disclosures enable investors to view the company's situation from perspectives other than its financial performance (Fatemi et al., 2018; Li et al., 2018; Mulchandani et al., 2022; Ng & Rezaee, 2020). The decrease in information asymmetry resulting from ESG disclosures improves transparency, reduces agency costs, strengthens investor confidence by reducing uncertainty, and can ultimately increase stock liquidity (R. Chen et al., 2023; Cui et al., 2018; Liang et al., 2023; Meng-tao et al., 2023; Wong & Zhang, 2021).

The relationship between ESG performance and corporate reputation can be explained by signaling theory. The information asymmetry results in outsiders interpreting any additional information beyond the mandatory financial reports as a signal to financial markets (Spence, 1973). A high ESG score is a signal that shows a high level of social and environmental responsibility and indicates that the business has a strategy that creates diversified integrated value and protects stakeholders. This positive signal sent to the markets not only attracts sensitive investors in terms of social responsibility and environmental awareness but also increases trust and support from all investors. The positive image and increased corporate reputation created in this context can influence investors' decisions in capital markets by enabling businesses to increase their financial performance and competitive advantages in the market (Wong et al., 2021; Wong & Zhang, 2021) and can increase stock liquidity by supporting stock trading (R. Chen et al., 2023; Luo, 2022).

ESG ratings also play a significant role in improving analyst forecast accuracy for companies and reducing firm risk (R. Chen et al., 2023; Meng-tao et al., 2023; Rossi & Harjoto, 2020). A good ESG performance indicates that the company has a solid corporate governance mechanism that attaches importance to continuous improvement, complies with relevant laws and regulations and is less likely to be penalized, has a higher sense of social responsibility, and has a high corporate reputation (Luo & Wu, 2022). Sassen et al. (2016) have demonstrated empirical evidence that businesses that provide non-financial information to the market with ESG disclosures can reduce firm risks, especially idiosyncratic risks. Investors are likely to turn to risky investments to obtain returns above the expected returns. However, past financial crises have shown that risk-taking behavior can cause great losses to investors. The risk-averse preferences of many investors, whether individual or corporate, are a factor that restricts them from investing in companies with a high-risk profile. Low firm risk has the potential to increase stock trading volume and liquidity to the extent that it raises investors' expectations about the company's future.

Institutional investors, as one of the main participants in capital markets, play a vital role in stock liquidity. ESG performance is an important dimension that provides a competitive advantage to businesses and provides data for measuring the sustainable development performance and long-term investment values of businesses. Institutional investors are not willing to invest in companies with low environmental and social performance (Dyck et al., 2019; Nofsinger et al., 2019). Therefore, high ESG performance has the potential to attract the attention of domestic and foreign institutional investors and encourage them to invest in these companies or increase their existing investments, which ultimately strengthens liquidity (R. Chen et al., 2023). Institutions have a scale advantage compared to individual investors in acquiring and processing information, with their professional and advanced investment knowledge and skills. Therefore, institutional investors are generally considered knowledgeable representatives. Co-investment behavior, which occurs as a result of the positive reaction of institutional investors to ESG performance being perceived by individual investors as a signal to invest in these companies, can pave the way for further increase in liquidity (Bai et al., 2022; R. Chen et al., 2023; Ding et al., 2017).

1.4 Hypothesis development

Based on the discussion above, two contrasting perspectives emerge. From the shareholder theory viewpoint, ESG activities can be seen as possible source of inefficiency or value destruction. In contrast, stakeholder theory suggests that ESG engagement enhances reputation, transparency, and investor confidence, thereby increasing market efficiency and liquidity. Accordingly, and consistent with the stakeholder-oriented perspective, we formulate our hypothesis as follows:

H1: Higher ESG scores enhance stock liquidity.

To the best of our knowledge, there is no prior study employing a comprehensive, large-scale, multi-country dataset and a robust methodology to test the ESG – liquidity nexus across Europe. This study thus provides a novel contribution to the literature by evaluating the impact of ESG performance on stock liquidity in European capital markets.

2 Research methodology

2.1 Data

European firms comprise the sample of this study. Both financial data and ESG data of firms listed in European stock exchanges were collected from the Refinitiv Eikon database for the period 2002 and 2022. Since ESG scores have been available since 2002, the sample period of the study starts in 2002. While some former papers (i.e., R. Chen et al., 2023; Meng-tao et al., 2023) obtained ESG scores from the Bloomberg database, we obtained ESG data for European listed firms from Refinitiv Eikon database (formerly known as Thomson Reuters Eikon) which is a highly utilized platform that provides financial data. It is one of the most frequently used databases for obtaining ESG scores, especially for European companies (i.e., Candio, 2024; Gawęda, 2024; Janicka & Sajnóg, 2023; Menicucci & Paolucci, 2025; Pinheiro et al., 2024). We started to form our sample by searching for European companies with ESG scores from the Refinitiv Eikon database between 2002 and 2022. This search yielded ESG scores for 2,188 firms from 25 countries (EU-27 excluding Bulgaria, Croatia, Estonia, Latvia, Lithuania, Malta, and including Iceland, Norway, Switzerland, and the United Kingdom) operating in 11 different industries according to TRBC economic sector classification given by Thomson Reuters. A total of 17,752 firm-year observations were yielded at the end of the first screening. Following previous studies (i.e., R. Chen et al., 2023; Meng-tao et al., 2023; Wang et al., 2023) firms operating in financial industries (TRBC code 55) were eliminated (3,178 firm-year observations) due to their unique reporting and regulatory requirements. The academic and educational services industry (TRBC code 63) was also eliminated (10 firm-year observations) due to its poor representation. Observations with missing data for variables in the baseline regression model were excluded and firm level continuous variables without upper/lower boundaries were winsorized by 1% from the bottom and top to control for the misleading effects of outliers. The final sample included 1,693 firms from nine industries and 12,052 firm-year observations. The distributions of the observations through countries, regions, industries, and years are shown in Tab. A1 in the appendix.

2.2 Model specification

Our baseline regression model for examining the relationship between ESG score and stock liquidity, which is exactly the same model used by R. Chen et al. (2023), Meng-tao et al. (2023), is specified in Equation (1). To the best of our knowledge, this approach has not been used for robustly testing the relationship between ESG score and stock liquidity across European Union states. We used the cluster(id) option to use robust standard errors to control for heteroskedasticity and autocorrelation.

$$\begin{aligned} &Stock\ Liquidity_{i,t} = c + \beta \times \\ &\times ESG\ Performance + \gamma \times \\ &\times Control\ Variables + \theta \times firmFE + \\ &+ \alpha \times yearFE + \varepsilon_{i,t} \end{aligned} \tag{1}$$

2.3 Variable description

Stock liquidity, which is the dependent variable of the study, is captured by the illiquidity measure of Amihud (2002) as in Equation (2). The illiquidity measure of Amihud (2002) has been reported as one of the best proxies among compared liquidity proxies (Fong et al., 2017; Marcelo & Quirós, 2006). Amihud (2002) defines illiquidity as the ratio of the absolute daily returns to the daily transaction volume in dollars averaged by the number of trading days. Since Equation (2) yields an illiquidity measure, following previous literature, the stock liquidity is calculated annually for each firm by the natural logarithm of illiquidity measure of Amihud (2002) (Equation (2)) multiplied by −1 so that it can be interpreted as the higher the values, the higher the stock liquidity. The illiquidity measure is calculated as follows:

$$ILLIQ_{i,y} = \frac{1}{D_{i,y}} \sum_{t=0}^{D_{i,y}} \frac{|R_{i,y,d}|}{VOL_{i,y,d}} \tag{2}$$

where: $|R_{i,y,d}|$ – the absolute daily stock returns (calculated by ln(price(*d*)/price(*d* − 1)) of stock *i* on day *d* of the year *y*; $Vol_{i,y,d}$ – the daily trading volumes (measured by transaction volumes in USD) of stock *i* on day *d* of year *y*; $D_{i,y}$ – the number of the trading days for stock *i* in year *y*.

As the measure for ESG performance, we used ESG scores by the Refinitiv Eikon database for the period 2002 and 2022. Apart from the data accessibility we chose the Refinitiv Eikon database as it covers over 90% of the global market cap (LSEG, 2023) in terms of ESG score which is the variable of interest of this study. We used ESG combined scores (ESG_COMB) as the main variable of interest as well as scores of ESG dimensions (ESG_E, ESG_S and ESG_G) where applicable.

Based on the previous literature we identified the following variables to control for the effects of firm level characteristics on stock liquidity to capture the effect of ESG more properly. The descriptives for the variables in our baseline regression are presented in Tab. 1.

Tab. 1: Descriptive statistics

	Mean	Median	Max.	Min.	Std. dev.	Obs.
LIQ	19.23	19.61	24.07	10.43	2.53	12,052
ESG_COMB	50.07	50.94	95.71	0.63	19.33	12,052
ROA	0.12	0.12	0.51	-0.30	0.09	12,052
LEV	0.26	0.25	0.75	0.00	0.15	12,052
SIZE	21.92	21.96	25.96	17.28	1.68	12,052
AGE	42.67	27.00	123.00	1.00	37.04	12,052
STR_INV	0.29	0.25	1.00	0.00	0.25	12,052
CASH	0.11	0.09	0.72	0.00	0.10	12,052

Source: own

The control variables are return on assets (ROA) which is calculated as the EBITDA over total assets, leverage (LEV) which is calculated as total debt over total assets, size (SIZE) which is proxied by the natural logarithm of total assets, age (AGE) measured by the natural logarithm of 2022 minus the incorporation year, strategic investors ratio (STR_INV) measured by shares held by strategic investors, and cash intensity (CASH) measured by cash and short term investments over total assets.

The descriptive statistics for the variables are shown in Tab. 1.

3 Results and discussion

3.1 Baseline regression results

The results of the baseline regression given by Equation (1) are shown in Tab. 2. Equation (1) is estimated with firm and year fixed effects. Regression results of the variable of interest, ESG combined score (ESG_COMB), and stock liquidity (LIQ) are shown without control variables in Model 1 and with control variables in Model 2. Models 3–5 show the results of sub-dimensions of combined ESG score (ESG_E, ESG_S, and ESG_G) and stock liquidity with control variables.

The baseline regression results reveal that higher ESG scores are positively associated with stock liquidity. Model 1 shows that higher ESG combined scores enhance stock liquidity alone, while Model 2 shows that this effect survives with the inclusion of control variables. This positive association shows that ESG endeavors have reflections in capital markets.

Besides the statistical significance of the estimated coefficients, which reveal the positive relationship, they suggest that ESG performance exerts a meaningful and economically non-negligible impact on firms' stock liquidity as well. Given in the variable description section that our dependent variable is defined

Tab. 2: Baseline regression results – Part 1

Dependent	Model 1	Model 2	Model 3	Model 4	Model 5
	LIQ	*LIQ*	*LIQ*	*LIQ*	*LIQ*
ESG_COMB	0.010***	0.004***			
	(6.028)	(4.157)			
ESG_E			0.003***		
			(3.661)		
ESG_S				0.002**	
				(1.662)	
ESG_G					0.001***
					(1.707)
ROA		2.369***	2.367***	2.367***	2.372***
		(10.243)	(10.168)	(10.142)	(10.166)
LEV		−1.388***	−1.399***	−1.386***	−1.383***
		(−7.683)	(−7.780)	(−7.659)	(−7.674)
SIZE		1.102***	1.103***	1.115***	1.116***
		(21.407)	(21.413)	(21.645)	(21.684)
AGE		−0.007	−0.006	−0.001	0.002
		(−0.119)	(−0.101)	(−0.013)	(0.025)
STR_INV		−1.780***	−1.783***	−1.785***	−1.772***
		(−9.375)	(−9.410)	(−9.400)	(−9.312)
CASH		0.674***	0.672***	0.675***	0.681***
		(2.940)	(2.937)	(2.936)	(2.948)

Tab. 2: Baseline regression results – Part 2

Dependent	Model 1	Model 2	Model 3	Model 4	Model 5
	LIQ	***LIQ***	***LIQ***	***LIQ***	***LIQ***
Cons	17.576***	−5.052***	−4.994***	−5.271***	−5.312***
	(147.690)	(−4.622)	(−4.557)	(−4.818)	(−4.863)
Year FE	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Observations	12,052	12,052	12,051	12,051	12,052
F	171.400	171.420	171.330	171.470	171.670
Adj. R^2	0.187	0.728	0.726	0.725	0.724

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

Source: own

as the negative natural logarithm of Amihud's (2002) illiquidity ratio ($LIQ = -\ln(ILLIQ_{Amihud})$), a positive coefficient indicates a reduction in illiquidity; hence an improvement in liquidity. The coefficient of 0.004 on the combined ESG score (Model 2) implies that a one-point increase in the ESG score (ranging from 0 to 100) leads, on average, to a 0.004 increase in $-\ln(ILLIQ)$. This corresponds approximately to a 0.4% decline in illiquidity, or equivalently, a 0.4% enhancement in stock liquidity. Therefore, a 10-point rise in a firm's ESG score (e.g., from 50 to 60) is associated with roughly a 4% improvement in liquidity, ceteris paribus. When decomposed into environmental, social, and governance pillars, the coefficients of 0.003, 0.002, and 0.001 respectively translate into 0.3, 0.2, and 0.1% reductions in illiquidity for each one-point increase. Here it should be noted that, since ESG scores are ordinal (rather than cardinal) in nature (Avramov, 2022), one should be cautious in operationalizing the economic interpretation of the results.

The association of control variables with stock liquidity is in line with the expectations. While ROA, SIZE, and CASH variables have statistically significant positive effect on stock liquidity, LEV and STR_INV variables have statistically significant negative effect.

The association of subdimensions of ESG with stock liquidity was tested by Model 3, Model 4, and Model 5. Each subdimension was found to affect stock liquidity positively. Judging by the economic significance (i.e., coefficient) and the statistical significance level (i.e., *t*-statistics) we found that the environmental subdimension of the combined ESG score has the highest effect, and the social subdimension of the combined ESG score has the lowest effect. According to our results, firms can expect positive effects of the disclosures they make regarding environmental, social, and governance issues on their stock liquidity, and prioritizing environmental disclosures among others would be more effective.

Besides, return on asset (ROA), size (SIZE), and cash intensity (CASH) were found to have a significantly positive effect, while leverage (LEV) and strategic investors ratio (STR_INV) were found to have a significantly negative effect on stock liquidity.

3.2 The effect of geographical region

Effects of regional characteristics on the relationship between ESG and stock liquidity are analyzed by subsample regressions. To explore the effect of geographical regions on the relationship between ESG and stock liquidity, we grouped European countries in our sample into four regions (namely Eastern Europe, Northern Europe, Southern Europe, and Western Europe) by using the United Nations geoscheme based on the M49 coding classification (United Nations, 2023). According to the UN geoscheme, Poland, Hungary, the Czech Republic, and Romania are considered as Eastern Europe. Western Europe includes Austria, Belgium, France, Germany, Luxembourg, the Netherlands, and Switzerland. Northern Europe is composed of Denmark, Finland, Iceland, the Republic of Ireland, Norway, Sweden, and the United Kingdom. Greece, Italy, Portugal, Slovenia, and Spain are classified as Southern Europe.

Regression results for the subsamples based on the geographical regions are presented in Tab. 3. While liquidity (LIQ) is the dependent variable for each model, ESG combined score

Tab. 3: Effect of ESG by region

Region	*ESG_COMB*	*ESG_E*	*ESG_S*	*ESG_G*	*Obs.*	Adj. R^2	*F*	Controls
Eastern	0.004				304	0.512	5.940	Included
	(0.663)							
		0.009**			304	0.521	6.230	Included
		(2.130)						
			0.006		304	0.520	6.010	Included
			(1.230)					
				-0.002	304	0.494	5.930	Included
				(-0.580)				
Northern	0.008***				5,680	0.740	89.370	Included
	(5.141)							
		0.005***			5,679	0.736	87.520	Included
		(3.796)						
			0.005***		5,680	0.737	89.900	Included
			(3.039)					
				0.001	5,680	0.733	88.100	Included
				(0.915)				
Southern	0.003				1,461	0.574	44.200	Included
	(0.842)							
		0.003			1,461	0.574	42.000	Included
		(1.150)						
			0.001		1,461	0.569	42.770	Included
			(0.127)					
				0.001	1,461	0.570	44.750	Included
				(0.508)				
Western	0.004***				4,607	0.723	83.300	Included
	(2.794)							
		0.004***			4,607	0.719	85.300	Included
		(2.819)						
			0.001		4,607	0.719	83.100	Included
			(0.936)					
				0.002**	4,607	0.721	83.250	Included
				(2.161)				

Source: own

(ESG_COMB) and subdimensions of ESG are the independent variables and the control variables defined earlier are also included. The results showed that while the coefficients of ESG_COMB were not statistically significant for Eastern and Southern European companies, they were significant for Northern and Western European companies. ESG combined scores were reflected in the stock liquidity for companies only operating in Northern and Western Europe. The effect of subdimensions of ESG on liquidity was not found homogeneous across geographical regions either. First, for none of the regions, every subdimension was found statistically significant. While environmental score (ESG_E) and social score (ESG_S) were found to have statistically positive effect on liquidity for Northern European companies, environmental score (ESG_E) and governance score (ESG_G) were found significant for Western European companies. While none of the subdimensions were found to have significant effect on liquidity for Southern European companies, only environmental score (ESG_E) was found to be significant for Eastern European companies. Out of the three subdimensions, environmental score (ESG_E) was found to be the most influential on stock liquidity.

3.3 Robustness tests

First, we checked for possible multicollinearity by tabulating cross-correlations of the variables. Cross-correlation coefficients are presented in Tab. 4. The correlation coefficients between independent variable pairs vary between 0.01 and 0.45, which do not indicate a series multicollinearity problem.

Tab. 4: Cross correlation of variables

	LIQ	*ESG_COMB*	*ROA*	*LEV*	*SIZE*	*AGE*	*STR_INV*	*CASH*
LIQ	1.00	0.42	0.14	0.01	0.77	0.02	−0.31	−0.10
ESG_COMB		1.00	0.00	0.09	0.45	0.12	−0.10	−0.07
ROA			1.00	−0.15	−0.09	−0.01	−0.02	0.01
LEV				1.00	0.18	−0.01	0.04	−0.25
SIZE					1.00	0.05	−0.07	−0.24
AGE						1.00	0.04	−0.05
STR_INV							1.00	0.08
CASH								1.00

Source: own

Second, the positive relationship between ESG and stock liquidity revealed by the analyses was also checked for possible endogeneity problems. Following Meng-tao et al. (2023), we estimated Equation (1) by using lagged independent variable of interest (ESG_COMB) or leaded dependent variable (LIQ). Control variables, described before, and firm and year fixed effects were included. Results presented in Tab. 5 show that both the economic and statistical significance and the direction of the coefficients persisted when the lagged or leaded variables were employed.

In order to add robustness to the results obtained by the baseline regressions, we estimated Equation (1) by using different parameters and checked whether the significant positive relationship between ESG and stock liquidity survived.

Differing from baseline regressions where firm and year fixed effects were included, Equation (1) was estimated by the inclusion of country fixed effects and year fixed effects (Model 1), by the inclusion of industry fixed effects and year fixed effects (Model 2), and Model 3 estimated the relationship by pooled OLS. To check the effect of the transition from local regulations to IFRS in 2005 on reporting standards, we ran the tests with data between 2005 and 2022 in Model 4. Finally, we used different control

Tab. 5: Lagging independent and leading dependent variable

	LIQ	*LIQ*	*LIQ* (*t* + 1)	*LIQ* (*t* + 2)
ESG_COMB			0.004***	0.004***
			(3.381)	(2.859)
ESG_COMB (*t* − 1)	0.004***			
	(3.501)			
ESG_COMB (*t* − 2)		0.003***		
		(2.871)		
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Controls	Included	Included	Included	Included
Observations	10,117	8,540	10,117	8,540
Firm FE	168.440	158.440	152.950	132.710
Adj. R^2	0.729	0.714	0.721	0.670

Note: *** significant at 1%.

Source: own

Tab. 6: Robustness checks – Part 1

Dependent	Panel A				Dependent	Panel B
	Model 1	Model 2	Model 3	Model 4		Model 5
	LIQ	*LIQ*	*LIQ*	*LIQ*		*LIQ*
ESG_COMB	0.01***	0.01***	0.01***	0.01***	*ESG_COMB*	0.01***
	(6.44)	(6.28)	(5.43)	(3.45)		(4.29)
ROA	2.51***	2.48***	5.52***	2.43***	*ROA2*	1.58***
	(11.33)	(11.17)	(16.95)	(10.63)		(6.96)
LEV	−1.48***	−1.41***	−1.34***	−1.33***	*LEV2*	−1.26***
	(−9.15)	(−8.65)	(−7.00)	(−7.57)		(−6.34)
SIZE	1.11***	1.13***	1.17***	1.10***	*SIZE2*	0.74***
	(40.21)	(41.59)	(61.40)	(20.44)		(9.77)
AGE	−0.08***	−0.03	−0.04	−0.02	*AGE*	0.01
	(−2.33)	(−0.82)	(−1.36)	(−0.36)		(0.05)
STR_INV	−2.07***	−2.03***	−2.52***	−1.86***	*STR_INV*	−2.00***
	(−13.30)	(−13.39)	(−18.87)	(−9.65)		(−9.83)
CASH	1.04***	0.95***	2.21***	0.55**	*CASH2*	0.01
	(5.02)	(4.54)	(8.18)	(2.49)		(−0.39)
CONS	−5.53***	−5.69***	−6.54***	0.00	*CONS*	3.53
	(−8.51)	(−9.53)	(−16.02)	(−3.39)		(2.22)

Tab. 6: Robustness checks – Part 2

Dependent	Panel A Model 1	Model 2	Model 3	Model 4	Dependent	Panel B Model 5
	LIQ	*LIQ*	*LIQ*	*LIQ*		*LIQ*
Firm FE	No	No	No	Yes	**Firm FE**	Yes
Year FE	Yes	Yes	No	Yes	**Year FE**	Yes
Country FE	Yes	No	No	No	**Country FE**	No
Industry FE	No	Yes	No	No	**Industry FE**	No
Observations	12,052	12,052	12,052	11,634	**Observations**	11,989
Wald chi²	2.33E+16	6,422.16			**Wald chi²**	
F			870.23	178.42	***F***	160.49
Adj. R^2	0.75	0.75	0.71	0.73	**Adj. R^2**	0.63

Source: own

variables instead of the ones used in baseline regression. Profitability was measured by return on average total assets (ROA2) instead of EBITDA over total assets (ROA). Leverage was calculated by total liabilities over total assets (LEV2) instead of total debt over total assets (LEV). Size was proxied by the natural logarithm of sales (SIZE2) instead of the natural logarithm of total assets (SIZE), and cash intensity was measured by the current ratio (CASH2) instead of cash and short-term investments over total assets. Tab. 6 shows the results of alternative models. The positive and significant effect of ESG on stock liquidity did not change.

3.4 Discussion

Our findings confirm the general consensus in the literature that ESG performance enhances stock liquidity. Similar to Aboud and Diab (2019), R. Chen et al. (2023), He et al. (2023), and Meng-tao et al. (2023), we find a positive and statistically significant relationship between ESG scores and stock liquidity. These studies argue that ESG activities reduce information asymmetry, improve corporate transparency, and attract institutional investors; the findings are also supported by our empirical results.

Our results closely align with R. Chen et al. (2023), who found that ESG performance increases liquidity by lowering agency costs and enhancing corporate reputation. Likewise, Wang et al. (2023) and Zhang et al. (2024) emphasized ESG's role in risk reduction and investor confidence, consistent with our finding that environmental disclosures, in particular, are positively associated with stock liquidity. Our baseline results, backed by robust tests support the notion that ESG activities are not merely ethical commitments but financially material signals that have influence on market efficiency. First, ESG disclosure (especially the governance pillar) can act as a mechanism to enhance the mitigation of some parts of the agency costs (e.g., monitoring costs) defined by Jensen and Meckling (1976). Second, as proposed by the information asymmetry theory, increased public information (hence reduced information asymmetry) can attract large investors because of increased liquidity of the securities (Diamond & Verrecchia, 1991). ESG disclosures can enhance transparency by reducing hidden information and consequently enhance stock liquidity. Firms that voluntarily communicate non-financial performance allow investors to evaluate management of the company beyond traditional financial indicators. This can decrease the uncertainty about firm value and facilitate the trading of the stocks. Third, high ESG scores (or even having ESG scores alone) can serve as credible market signals of management quality as implied by the signaling theory (Certo, 2003; Spence, 1973). Investors can interpret high ESG scores as indicators of sound internal governance, which strengthens confidence and fosters active trading. Last but not least, the rapidly growing body of theoretical and empirical research on social preferences of investors (e.g., Bauer et al., 2021; Riedl & Smeets, 2017) also suggests that high ESG scores attract sustainability-oriented investors and hence directly enhance stock liquidity.

However, our study diverges from prior research in the relative importance of ESG subdimensions. While Meng-tao et al. (2023) identified governance as the most influential and environmental as the least impactful ESG dimension, our results reveal the opposite: the environmental component exerts the strongest impact, followed by governance and social dimensions. This pattern can be conceptually explained by differences in visibility, measurability, and investor salience. Environmental performance, such as emissions reduction, resource efficiency, and renewable energy use, provides quantifiable and externally verifiable information that sends strong signals of regulatory compliance and operational efficiency. In the European context, where the EU Green Deal and carbon neutrality targets dominate policy discourse, environmental achievements have become particularly salient to investors. In contrast, social performance (e.g., employee welfare or community engagement) often yields longer-term reputational benefits that are less immediately priced by capital markets. This discrepancy may also reflect regional differences in investor priorities. In the European context, especially Northern and Western regions, climate-related issues are more deeply embedded in policy and investment practices, potentially increasing the importance of environmental disclosures.

Our regional analysis adds an important nuance not widely covered in earlier studies. While prior works such as Subramaniam et al. (2016) and Wong and Zhang (2021) provide evidence of a uniform positive relationship between ESG and liquidity, our study highlights that this relationship is not homogeneous across Europe. Specifically, we found that ESG scores are significantly associated with stock liquidity in Northern and Western Europe, but this association is not statistically significant in Southern and Eastern Europe. This regional disparity has not been thoroughly addressed in previous literature, representing a key contribution of our work. Apart from the sample selection constraints this divergence can be better understood and further explored from the viewpoints of some theories. This phenomenon can be explained by varieties of capitalism literature stating that coordination of firms' activities and firms' responses to other actors differ across liberal market economies, coordinated market economies (Hall & Soskice, 2003) and mixed market economies (Hassel, 2014). Institutional complementarities in different European countries, which shape varieties of capitalism across European countries, can also help to understand this divergence (Hall & Soskice, 2003). Organizational legitimacy framework (Suchman, 1995) can also provide an alternative explanation for why sustainability related issues are not equally rewarded in different European stock markets in Northern and Western Europe awareness about sustainability issues can be higher and sustainability regulations such as the EU Non-Financial Reporting Directive (2014/95/EU) and the Sustainable Finance Disclosure Regulation (2021) (European Parliament and Council of the European Union, 2014, 2021) can be more deeply adopted and institutionalized. This can trigger normative pressures for enhanced ESG practices. As a result, firms in these regions experience stronger legitimacy gains and market trust when they perform well on ESG metrics. In contrast, the social and regulatory expectations around sustainability can be weaker in Southern and Eastern Europe; therefore, ESG performance provides a less potent legitimacy signal to investors. Lower enforcement of disclosure standards and smaller sustainability-oriented investor bases reduce the informational and signaling value of ESG data. Consequently, an equivalent improvement in ESG performance leads to a more muted investor response and generates weaker liquidity effects.

While Fatemi et al. (2018) and Saygili et al. (2022) raised concerns about negative or ambiguous ESG effects, our results do not support this view. All three ESG dimensions show positive and significant associations with liquidity. Furthermore, in contrast to Capelle-Blancard and Petit (2019), who reported that markets penalize poor ESG performance but do not reward strong performance, our findings show that high ESG scores are positively recognized, particularly in developed regions.

In summary, our study confirms the positive impact of ESG performance on stock liquidity, while also contributing new evidence of regional disparities and a greater emphasis on environmental factors within European markets.

Conclusions

In this study, the impact of corporate ESG performance on stock liquidity is examined. Our research findings show that this effect is positive, that is, to the extent that firms' ESG performance increases, their performance in capital markets and stock liquidity also increases. The mechanism of this relationship can be briefly summarized as follows. As ESG performance increases, there will likely be a decrease in information asymmetry and corporate risks and an increase in corporate reputation. In addition, institutional investors' investment preferences in favor of high-scored firms result in both these investors carrying out large transaction volumes and individual investors perceiving these transactions as signals and directing their investments to these firms.

The result indicating that the association of ESG performance and stock liquidity differs across European regions is one of the interesting findings of this research, which needs further attention by future studies. Our results show that investors in Eastern and Southern Europe do not attach as much importance to sustainability issues as those in Western and Northern Europe. The fact that Western and Northern European countries have larger economies compared to Eastern and Southern European ones causes the GDP per capita and per capita income to be higher in the West and North (Eurostat, 2023). This situation may result in investors, whether they have environmentally friendly attitudes and values or care deeply about human rights and social equality, directing their relatively limited savings to investment instruments focused on profit maximization, instead of environmentally friendly or socially responsible financial products. There are also differences across Europe in terms of government initiatives supporting sustainable investments. While countries that share the characteristics of the Scandinavian socio-economic model (Denmark, Finland, the Netherlands, and Sweden) and countries with the Anglo-Saxon socio-economic model (England and Ireland) are much more mature in sustainable investing, Mediterranean and Transition countries are further behind in this regard. Our results support the findings of Steurer et al. (2008), which indicate that there is a West-East divide as well as a North-South divide in terms of ESG policies in Europe.

Putting this discrepancy aside, our findings show that investor behavior, which in the past was largely focused on profitability, has transformed and is influenced by the activities carried out by companies in the environmental and social fields and good governance practices. Investors today prefer more sustainable companies in the context of nature protection, social justice, and good governance.

Additionally, our findings show that firms that incorporate ESG into their processes can improve their reputation, reduce their risks, and potentially attract more investors. Investors, as critical stakeholders, have a significant impact on the strategies firms implement and the decisions they make. In this context, investors' support and trust in ESG legitimizes firms' ESG practices. Ignoring sustainability and ESG, which have become one of the most important strategies of all modern organizations around the world in order to maintain competitiveness and long-term success, would be a big strategic mistake.

Like any study, this research has its limitations. Among various ESG information providers, such as Bloomberg, MSCI/KLD, Fortune, and the Dow Jones Sustainability Indices (DJSI), we relied on ESG scores from Refinitiv Eikon due to data availability. Future studies could expand the sample by incorporating these additional ESG sources. Another limitation stems from the regional imbalance in company representation; our conclusion – that investors in Western and Northern Europe prioritize sustainability more than their counterparts in Eastern and Southern Europe – is based on an uneven distribution of companies across these regions. The unequal representation of regions weakens the generalizability of the results. Expanding the sample to include more companies from Eastern and Southern Europe could help address this issue. Besides, the regional differences deserve further research of the relationship between ESG performance and stock liquidity in the context of economic development, cultural structure and government initiatives. Another economic and methodological limitation of our analysis lies in the potential presence of reverse causality between ESG performance and stock liquidity. Firms with higher stock liquidity may have greater visibility and better images in capital markets, which can motivate enhanced ESG practices and disclosure. More liquid stocks are more likely to be traded and held heavily by institutional investors and are more likely to face closer public scrutiny, which would also enhance ESG practices by these firms. Thus, the positive association observed between ESG scores and stock liquidity may partially reflect a bidirectional relationship rather than a purely causal effect from ESG to liquidity. While firm and year fixed effects were included in our models to account for unobserved heterogeneity, the reverse causality could not be ruled out in our baseline regressions. Although the lagging independent and leading dependent variable in the robustness tests were attempts to mitigate the endogeneity problems (including reverse causality), future research could address this issue by employing more specific tests and methods to more robustly identify the causal direction. Last but not least, the relationship between ESG performance and stock liquidity is an area open to further investigation. The dependent variable of this study, stock liquidity, was proxied by the illiquidity measure of Amihud (2002). Alternative measures of stock liquidity can be used by future research to test the proposed relationships.

References

Aboud, A., & Diab, A. (2019). The financial and market consequences of environmental, social and governance ratings: The implications of recent political volatility in Egypt. *Sustainability Accounting, Management and Policy Journal*, *10*(3), 498–520. https://doi.org/10.1108/SAMPJ-06-2018-0167

Agarwala, N., Jana, S., & Sahu, T. N. (2024). ESG disclosures and corporate performance: A non-linear and disaggregated approach. *Journal of Cleaner Production*, *437*, 140517. https://doi.org/10.1016/j.jclepro.2023.140517

Alareeni, B. A., & Hamdan, A. (2020). ESG impact on performance of US S&P 500-listed firms. *Corporate Governance*, *20*(7), 1409–1428. https://doi.org/10.1108/CG-06-2020-0258

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56. https://doi.org/10.1016/S1386-4181(01)00024-6

Amihud, Y., Mendelson, H., & Pedersen, L. H. (2006). Liquidity and asset prices. *Foundations and Trends in Finance*, *1*(4), 269–364. https://doi.org/10.1561/0500000003

Atan, R., Alam, M. M., Said, J., & Zamri, M. (2018). The impacts of environmental, social, and governance factors on firm performance: Panel study of Malaysian companies. *Management of Environmental Quality*, *29*(2), 182–194. https://doi.org/10.1108/MEQ-03-2017-0033

Avramov, D., Cheng, S., Lioui, A., & Tarelli, A. (2022). Sustainable investing with ESG rating uncertainty. *Journal of Financial Economics*, *145*(2), 642–664. https://doi.org/10.1016/j.jfineco.2021.09.009

Aydoğmuş, M., Gülay, G., & Ergun, K. (2022). Impact of ESG performance on firm value and profitability. *Borsa Istanbul Review*, *22*(2), 119–127. https://doi.org/10.1016/j.bir.2022.11.006

Bagh, T., Fuwei, J., & Khan, M. A. (2024). Corporate ESG investments and firm's value under the real-option framework: Evidence from two world-leading economies. *Borsa Istanbul Review*, *24*(2), 324–340. https://doi.org/10.1016/j.bir.2024.01.002

Bai, X., Han, J., Ma, Y., & Zhang, W. (2022). ESG performance, institutional investors' preference and financing constraints: Empirical evidence from China. *Borsa Istanbul Review*, *22*(2), 157–168. https://doi.org/10.1016/j.bir.2022.11.013

Bauer, R., Ruof, T., & Smeets, P. (2021). Get real! Individuals prefer more sustainable investments. *The Review of Financial Studies*, *34*, 3976–4043. https://doi.org/10.1093/rfs/hhab037

Bhowmick, D., Salunkhe, H. A., & Boralkar, M. (2025). The impact of environmental, social, and governance (ESG) objectives' effect on investment decisions: Retail investors' knowledge and views on sector-specific ESG Factors. In A. Sharma, M. George, P. Pandey, V. Joshi, & M. Sharma (Eds), *The synergistic effect of sustainable business practices on corporate performance* (pp. 61–75). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-3375-3_4

Bihari, A., Dash, M., Muduli, K., Kumar, A., Luthra, S., & Upadhyay, A. (2025). Sustainable development investment decision: Do environmental, social and governance (ESG) and behavioral biases factors matter? *Sustainable Development*, *33*(S1), 1161–1181. https://doi.org/10.1002/sd.70056

Brammer, S., Brooks, C., & Pavelin, S. (2006). Corporate social performance and stock returns: UK evidence from disaggregate measures. *Financial Management*, *35*(3), 97–116. https://doi.org/10.1111/j.1755-053X.2006.tb00149.x

Broadstock, D. C., Chan, K., Cheng, L. T., & Wang, X. (2021). The role of ESG performance during times of financial crisis: Evidence from COVID-19. *Finance Research Letters*, *38*, 101716. https://doi.org/10.1016/j.frl.2020.101716

Candio, P. (2024). The effect of ESG and CSR attitude on financial performance in Europe: A quantitative re-examination. *Journal of Environmental Management*, *354*, 120390. https://doi.org/10.1016/j.jenvman.2024.120390

Capelle-Blancard, G., & Petit, A. (2019). Every little helps? ESG news and stock market reaction. *Journal of Business Ethics*, *157*, 543–565. https://doi.org/10.1007/s10551-017-3667-3

Certo, T. (2003). Influencing initial public offering investors with prestige: Signaling with board structures. *Academy of Management Review*, *28*(3), 432–446. https://doi.org/10.5465/amr.2003.10196754

Chen, R., Liu, Y., Jiang, Y., & Liu, J. (2023). Does ESG performance promote vitality of capital market? Analysis from the perspective of stock liquidity. *Frontiers in Environmental Science*, *11*, 1132845. https://doi.org/10.3389/fenvs.2023.1132845

Chen, S., Song, Y., & Gao, P. (2023). Environmental, social, and governance (ESG) performance and financial outcomes: Analyzing the impact of ESG on financial performance. *Journal of Environmental Management*, *345*, 118829. https://doi.org/10.1016/j.jenvman.2023.118829

Chen, Y. C., Hung, M., & Wang, Y. (2018). The effect of mandatory CSR disclosure on firm profitability and social externalities: Evidence from China. *Journal of Accounting and Economics*, *65*(1), 169–190. https://doi.org/10.1016/j.jacceco.2017.11.009

Chen, Z., & Xie, G. (2022). ESG disclosure and financial performance: Moderating role of ESG investors. *International Review of Financial Analysis*, *83*, 102291. https://doi.org/10.1016/j.irfa.2022.102291

Cui, J., Jo, H., & Na, H. (2018). Does corporate social responsibility affect information asymmetry? *Journal of Business Ethics*, *148*(3), 549–572. https://doi.org/10.1007/s10551-015-3003-8

Diamond, D. W., & Verrecchia, R. E. (1991). Disclosure, liquidity, and the cost of capital. *The Journal of Finance*, *46*(4), 1325–1359. https://doi.org/10.1111/j.1540-6261.1991.tb04620.x

Ding, M., Nilsson, B., & Suardi, S. (2017). Foreign institutional investment, ownership, and liquidity: Real and informational frictions. *Financial Review*, *52*(1), 101–144. https://doi.org/10.1111/fire.12126

Duque-Grisales, E., & Aguilera-Caracuel, J. (2021). Environmental, social and governance (ESG) scores and financial performance of multilatinas: Moderating effects of geographic international diversification and financial slack. *Journal of Business Ethics*, *168*, 315–334. https://doi.org/10.1007/s10551-019-04177-w

Dyck, A., Lins, K. V., Roth, L., & Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics*, *131*(3), 693–714. https://doi.org/10.1016/j.jfineco.2018.08.013

European Parliament and Council of the European Union. (2014). *Directive (2014/95/EU) of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU – Non-financial Reporting Directive.*

European Parliament and Council of the European Union. (2021). *Regulation (EU) 2019/2088 of the European Parliament and of the Council of 27 November 2019 on sustainability-related disclosures in the financial services sector.*

Eurostat. (2023). *GDP per capita, consumption per capita and price level indices*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=GDP_per_capita,_consumption_per_capita_and_price_level_indices

Fatemi, A., Fooladi, I., & Tehranian, H. (2015). Valuation effects of corporate social responsibility. *Journal of Banking & Finance*, *59*, 182–192. https://doi.org/10.1016/j.jbankfin.2015.04.028

Fatemi, A., Glaum, M., & Kaiser, S. (2018). ESG performance and firm value: The moderating role of disclosure. *Global Finance Journal*, *38*, 45–64. https://doi.org/10.1016/j.gfj.2017.03.001

Fong, K. Y. L., Holden, C. W., & Trzcinka, C. A. (2017). What are the best liquidity proxies for global research? *Review of Finance*, *21*(4), 1355–1401. https://doi.org/10.1093/rof/rfx003

Freeman, R. E. (1984). *Strategic management: A stakeholder approach*. Pitman Publishing.

Friedman, M. (1970). The social responsibility of business is to increase its profits.

The New York Times Magazine. https://www.nytimes.com/1970/09/13/archives/a-friedman-doctrine-the-social-responsibility-of-business-is-to.html

Gawęda, A. (2024). Does the sustainability of the country differentiate the ESG of companies and how it affects the relationship between ESG and firm value? Evidence from the European Union. *Annales Universitatis Mariae Curie-Skłodowska, Sectio H Oeconomia*, *58*(4), 7–23. https://doi.org/10.17951/h.2024.58.4.7-23

Hall, P. A., & Soskice, D. (2003). Varieties of capitalism and institutional complementarities. In R. Franzese, P. Mooslechner, & M. Schürz (Eds), *Institutional conflicts and complementarities* (pp. 43–76). Springer US. https://doi.org/10.1007/978-1-4757-4062-2_3

Hassel, A. (2014). *Adjustments in the Eurozone: Varieties of capitalism and the crisis in Southern Europe* [LEQS paper]. London School of Economics and Political Science (LSE), European Institute.

He, F., Feng, Y., & Hao, J. (2023). Corporate ESG rating and stock market liquidity: Evidence from China. *Economic Modelling*, *129*, 106511. https://doi.org/10.1016/j.econmod.2023.106511

Horvathova, E. (2010). Does environmental performance affect financial performance? A meta-analysis. *Ecological Economics*, *70*(1), 52–59. https://doi.org/10.1016/j.ecolecon.2010.04.004

Hutchings, G. (2025). The impact of an ESG nudge on retail investor engagement in sustainable finance: Evidence from an investment decision experiment. *Sustainability Accounting, Management and Policy Journal*, *16*(6), 1807–1830. https://doi.org/10.1108/sampj-10-2024-1104

Janicka, M., & Sajnóg, A. (2023). Do environmental and economic performance go hand in hand? An industrial analysis of European Union companies with the non-parametric data envelopment analysis method. *Corporate Social Responsibility and Environmental Management*, *30*(5), 2590–2605. https://doi.org/10.1002/csr.2504

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, *3*(4), 305–360. https://doi.org/10.1016/0304-405X(76)90026-X

Kao, F. C. (2023). How do ESG activities affect corporate performance? *Managerial and Decision Economics*, *44*(7), 4099–4116. https://doi.org/10.1002/mde.3944

Khan, U., & Liu, W. (2023). The link between green innovations, corporate performance, ESG activities, and sharing economy. *Environmental Science and Pollution Research*, *30*(32), 78763–78775. https://doi.org/10.1007/s11356-023-27722-7

Krüger, P. (2015). Corporate goodness and shareholder wealth. *Journal of Financial Economics*, *115*(2), 304–329. https://doi.org/10.1016/j.jfineco.2014.09.008

Li, Y., Gong, M., Zhang, X.-Y., & Koh, L. (2018). The impact of environmental, social, and governance disclosure on firm value: The role of CEO power. *The British Accounting Review*, *50*(1), 60–75. https://doi.org/10.1016/j.bar.2017.09.007

Liang, Y., Xue, C., & Zhang, J. (2023). The impact of ESG ratings on stock liquidity risk: Evidence from the Chinese market. *Review of Integrative Business and Economics Research*, *12*(4), 1–16. https://doi.org/10.58745/riber_12-4_1-16

Lins, K. V., Servaes, H., & Tamayo, A. (2017). Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. *The Journal of Finance*, *72*(4), 1785–1824. https://doi.org/10.1111/jofi.12505

Liu, P., Zhu, B., Yang, M., & Chu, X. (2022). ESG and financial performance: A qualitative comparative analysis in China's new energy companies. *Journal of Cleaner Production*, *379*, 134721. https://doi.org/10.1016/j.jclepro.2022.134721

LSEG. (2023). *Environmental, social and governance scores from LSEG*. https://www.lseg.com/content/dam/data-analytics/en_us/documents/methodology/lseg-esg-scores-methodology.pdf

Luo, D. (2022). ESG, liquidity, and stock returns. *Journal of International Financial Markets, Institutions and Money*, *78*, 101526. https://doi.org/10.1016/j.intfin.2022.101526

Luo, K., & Wu, S. (2022). Corporate sustainability and analysts' earnings forecast accuracy: Evidence from environmental, social and governance ratings. *Corporate Social Responsibility and Environmental Management*, *29*(5), 1465–1481. https://doi.org/10.1002/csr.2284

Malik, N., & Kashiramka, S. (2024). Impact of ESG disclosure on firm performance and cost of debt: Empirical evidence from India. *Journal* *of Cleaner Production*, *448*, 141582. https://doi.org/10.1016/j.jclepro.2024.141582

Marcelo, J. L. M., & Quirós, M. del M. M. (2006). The role of an illiquidity risk factor in asset pricing: Empirical evidence from the Spanish stock market. *The Quarterly Review of Economics and Finance*, *46*(2), 254–267. https://doi.org/10.1016/j.qref.2005.08.002

Margolis, J. D., Elfenbein, H. A., & Walsh, J. P. (2009). Does it pay to be good and does it matter? A meta-analysis of the relationship between corporate social and financial performance. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.1866371

Meng-tao, C., Da-Peng, Y., Wei-Qi, Z., & Qi-Jun, W. (2023). How does ESG disclosure improve stock liquidity for enterprises – Empirical evidence from China. *Environmental Impact Assessment Review*, *98*, 106926. https://doi.org/10.1016/j.eiar.2022.106926

Menicucci, E., & Paolucci, G. (2025). Board characteristics and ESG performance. Evidence from the European utilities sector. *International Journal of Energy Sector Management*, *20*(3), 786–813. https://doi.org/10.1108/IJESM-03-2025-0002

Miralles-Quirós, M., Miralles-Quirós, J., & Valente Gonçalves, L. (2018). The value relevance of environmental, social, and governance performance: The Brazilian case. *Sustainability*, *10*(3), 574. https://doi.org/10.3390/su10030574

Mohammad, W. M. W., & Wasiuzzaman, S. (2021). Environmental, social and governance (ESG) disclosure, competitive advantage and performance of firms in Malaysia. *Cleaner Environmental Systems*, *2*, 100015. https://doi.org/10.1016/j.cesys.2021.100015

Ng, A. C., & Rezaee, Z. (2020). Business sustainability factors and stock price informativeness. *Journal of Corporate Finance*, *64*, 101688. https://doi.org/10.1016/j.jcorpfin.2020.101688

Nofsinger, J. R., Sulaeman, J., & Varma, A. (2019). Institutional investors and corporate social responsibility. *Journal of Corporate Finance*, *58*, 700–725. https://doi.org/10.1016/j.jcorpfin.2019.07.012

Noja, G., G., Baditoiu, B. R., Buglea, A., Munteanu, V. P., & Gligor Cimpoieru, D. C. (2024). The impact of environmental, social and governance policies on companies' financial and economic performance: A comprehensive approach and new empirical evidence. *E&M Economics and Management*, *27*(1), 121–144.https://doi.org/10.15240/tul/001/2024-1-008

Pinheiro, A. B., dos Santos, J. I. A. S., Cherobim, A. P. M. S., & Segatto, A. P. (2024). What drives environmental, social and governance (ESG) performance? The role of institutional quality. *Management of Environmental Quality: An International Journal*, *35*(2), 427–444. https://doi.org/10.1108/MEQ-03-2023-0091

Riedl, A., & Smeets, P. (2017). Why do investors hold socially responsible mutual funds? *The Journal of Finance*, *72*(6), 2505–2549. https://doi.org/10.1111/jofi.12547

Rossi, F., & Harjoto, M. A. (2020). Corporate non-financial disclosure, firm value, risk, and agency costs: Evidence from Italian listed companies. *Review of Managerial Science*, *14*(5), 1149–1181. https://doi.org/10.1007/s11846-019-00358-z

Sanseverino, A., González-Ramírez, J., & Cwik, K. (2024). Do ESG progress disclosures influence investment decisions? *International Journal of Disclosure and Governance*, *21*(1), 107–126. https://doi.org/10.1057/s41310-023-00198-0

Sassen, R., Hinze, A.-K., & Hardeck, I. (2016). Impact of ESG factors on firm risk in Europe. *Journal of Business Economics*, *86*(8), 867–904. https://doi.org/10.1007/s11573-016-0819-3

Saygili, E., Arslan, S., & Birkan, A. O. (2022). ESG practices and corporate financial performance: Evidence from Borsa Istanbul. *Borsa Istanbul Review*, *22*(3), 525–533. https://doi.org/10.1016/j.bir.2021.07.001

Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, *87*(3), 355–374. https://doi.org/10.2307/1882010

Steurer, R., Margula, S., & Martinuzzi, A. (2008). *Socially responsible investment in EU member states: Overview of government initiatives and SRI experts' expectations towards governments* [Final report to the EU High-Level Group on CSR]. Vienna University of Economics and Business Administration.

Subramaniam, R. K., Samuel, S. D., & Mahenthiran, S. (2016). Liquidity implications of corporate social responsibility disclosures: Malaysian evidence. *Journal of International Accounting Research*, *15*(1), 133–153. https://doi.org/10.2308/jiar-51204

Suchman, M. C. (1995). Managing legitimacy: Strategic and institutional approaches.

The Academy of Management Review, *20*(3), 571–610. https://doi.org/10.2307/258788

United Nations. (2023). *Standard country or area codes for statistical use (M49)*. UN Department of Economic and Social Affairs, Statistics Division. https://unstats.un.org/unsd/methodology/m49/

Wang, K., Li, T., San, Z., & Gao, H. (2023). How does corporate ESG performance affect stock liquidity? Evidence from China. *Pacific-Basin Finance Journal*, *80*, 102087. https://doi.org/10.1016/j.pacfin.2023.102087

Wong, J. B., & Zhang, Q. (2021). Stock market reactions to adverse ESG disclosure via media channels. *The British Accounting Review*, *54*(1), 101045. https://doi.org/10.1016/j.bar.2021.101045

Wong, W. C., Batten, J. A., Ahmad, A. H., Mohamed-Arshad, S. B., Nordin, S., & Adzis, A. A. (2021). Does ESG certification add firm value? *Finance Research Letters*, *39*, 101593. https://doi.org/10.1016/j.frl.2020.101593

Ye, C., Song, X., & Liang, Y. (2022). Corporate sustainability performance, stock returns, and ESG indicators: Fresh insights from EU member states. *Environmental Science and Pollution Research*, *29*, 87680–87691. https://doi.org/10.1007/s11356-022-20789-8

Yoo, S., & Managi, S. (2022). Disclosure or action: Evaluating ESG behavior towards financial performance. *Finance Research Letters*, *44*, 102108. https://doi.org/10.1016/j.frl.2021.102108

Zhang, C., Hao, D., Gao, L., Xia, F., & Zhang, L. (2024). Do ESG ratings improve capital market trading activities? *International Review of Economics & Finance*, *93*(A), 195–210. https://doi.org/10.1016/j.iref.2024.03.027

Appendix

Tab. A1: Distribution of observations among countries and industries

Panel A

	Basic materials	Consumer cyclicals	consumer non-cyclicals	Energy	Healthcare	Industrials	Real estate	Technology	Utilities	Total
Austria	5			20		61	29		4	**119**
Belgium	56	25	18	23	59	46	35	56	15	**333**
Czech Republic									15	**15**
Denmark	43	39	48	28	98	130		39	7	**432**
Finland	91	104	51	17	22	144	11	76	17	**533**
France	90	427	125	67	114	409	102	172	34	**1,540**
Germany	233	218	50	18	155	291	72	288	54	**1,379**
Greece	2	17	4	37		79	4	19	33	**195**
Hungary		3	4	14	10			14		**45**
Iceland			5	2		6		3		**16**
Ireland; Rep.	11	30	35		4		5		2	**87**
Italy	23	132	22	72	28	127	4	65	109	**582**
Luxembourg		6								**6**
Netherlands	89	15	38	59	25	89	29	63		**407**
Norway	70	46	67	158		79	12	52	9	**493**
Poland	40	3	33	28		28	11	47	46	**236**
Portugal	21		38	15		9		32	37	**152**
Romania		1		5	1				1	**8**
Slovenia					4			1		**5**
Spain	48	74	31	34	52	139	34	53	62	**527**
Sweden	164	186	48	20	123	358	109	249	22	**1,279**
Switzerland	123	103	60		92	218	72	145	10	**823**
United Kingdom	369	752	197	168	180	703	229	186	56	**2,840**
Total	**1,478**	**2,181**	**874**	**785**	**967**	**2,916**	**758**	**1,560**	**533**	**12,052**

Panel B

	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Total
Eastern							1	7	10	16	16	19	19	20	20	21	22	29	31	35	38	**304**
Northern	39	61	90	128	139	148	156	169	174	179	190	193	202	224	246	263	360	453	674	790	802	**5,680**
Southern	7	19	20	24	31	37	44	46	48	51	54	55	55	61	62	72	104	144	152	183	192	**1,461**
Western	45	67	70	98	113	122	140	151	155	168	173	181	188	199	200	218	327	416	479	542	555	**4,607**
Total	**91**	**147**	**180**	**250**	**283**	**307**	**341**	**373**	**387**	**414**	**433**	**448**	**464**	**504**	**528**	**574**	**813**	**1,042**	**1,336**	**1,550**	**1,587**	**12,052**

Source: own

Explainable credit risk modeling with hybrid tabular deep learning and adaptive feature routing

R. John Martin[1], *Ahmad Khawaji*[2], *S. Sujatha*[3]

[1] *Jazan University, College of Engineering and Computer Science, Department of Computer Science, Saudi Arabia, ORCID: 0000-0003-4997-0203, jmartin@jazanu.edu.sa (corresponding author);*
[2] *Jazan University, College of Engineering and Computer Science, Department of Computer Science, Saudi Arabia, ORCID: 0009-0001-6380-8563, ahkhawaji@jazanu.edu.sa;*
[3] *Bharathi Women's College, Department of Computer Science, India, ORCID: 0009-0007-4900-2567, sujaphd@gmail.com.*

Abstract: *Well defined and transparent credit risk evaluation continues to be a fundamental issue in financial decision-making, particularly when dealing with intricate borrower profiles in the forms of organized tabular data. While traditional machine learning models often lack interpretability for predictive accuracy, recent deep learning approaches struggle to generalize across such data formats. We propose an innovative hybrid Tabular Deep Learning system that combines feature tokenizer transformer (FT-transformer) and TabNet architectures with an adaptive feature routing (AFR) method. The AFR module dynamically identifies significant aspects for each data instance, facilitating context-aware representation learning and enhancing generalization across other borrower groups. To guarantee explainability and compliance with regulations, our approach integrates a multimodal interpretability package that includes Shapley additive explanations (SHAP)-based attributions, attention mapping, and counterfactual reasoning for actionable what-if analysis. Comprehensive experiments on the benchmark dataset reveal the model's exceptional performance, attaining an AUC of 0.985, an F1-score of 0.972, and a premier ranking according to the Gini coefficient. Visual metrics of AUC vs. Gini coefficient, cost-benefit curves, and the counterfactual dashboards showcase the model's transparency and practical applicability. This study observes the application of explainable AI in credit risk modeling by successfully reconciling the balance between higher predictive accuracy and interpretability, hence facilitating explainable financial decision-making scenarios.*

Keywords: *Credit risk modeling, deep learning, explainable AI, TabNet, decision-making.*

JEL Classification: *G17, G21, G32, G51, H63.*

APA Style Citation: Martin, R. J., Khawaji, A. & Sujatha, S. (2026). Explainable credit risk modeling with hybrid tabular deep learning and adaptive feature routing. *E&M Economics and Management*, *29*(2), 197–219. https://doi.org/10.15240/tul/001/2026-5-005

Early Access Publication Date: *March 31, 2026.*

Introduction

In banking, reliable credit risk analysis is essential to stable lending practices and the resilience of the financial system. It facilitates financial institutions to assess the probability of borrower default and to make informed credit decisions that safeguard capital, manage financial risk, and adhere to regulatory requirements. Considering the magnitude and complexity of personal and corporate lending markets, precise and responsible credit risk modeling is not merely a predictive task, it is vital for protecting portfolio integrity, reducing non-performing loans, and upholding economic stability.

Historically, credit risk assessment has depended on statistical models like logistic regression and decision trees, which provide clarity and simplicity in application. Nonetheless, these models frequently fail to accurately represent the nonlinear, high-dimensional interactions inherent in real-world financial statistics, particularly as borrower behavior becomes more complex. Consequently, financial institutions have adopted sophisticated AI and machine learning (ML) methodologies, which have shown significant potential in uncovering hidden patterns and enhancing forecasting accuracy (Chang et al., 2024; Jeribi et al., 2024). Ensemble approaches such as extreme gradient boosting (XGBoost) and light gradient boosting machine (LightGBM) have been prominent for their efficacy in managing extensive credit datasets. Although these models have enhanced forecasting accuracy, their unclear decision-making processes provide difficulties for regulatory compliance, risk management, and user confidence, all of which are essential in the financial sector. Beyond performance, the importance of transparency led to a growing body of research in explainable AI (XAI), focused on improving the interpretability of credit risk models while maintaining accuracy (Hollmann et al., 2025; Martin et al., 2024). Roy and Vasa (2025) highlight that a significant obstacle to the practical implementation of AI-driven financial models is the absence of explainability and user trust, notwithstanding their shown predictive capabilities. Consequently, creating methodologies that are both precise and transparent, as well as auditable, has emerged as a primary goal in contemporary credit risk analytics.

At the same time, tabular deep learning (TDL), which is the application of deep learning to structured data, has been gaining popularity (Ye et al., 2024). Although deep learning has historically excelled in areas such as image and text processing (Martin et al., 2023), contemporary architectures like TabNet (Arik & Pfister, 2021), TabTransformer (Huang et al., 2020), and Tabular Prediction Foundation Network (TabPFN) (Liu & Ye, 2025) have been explicitly engineered for tabular datasets (for further explanation of the terms, see Tab. A1 in appendix). The TabPFN, a few-shot transformer-based learner, has demonstrated impressive performance in low-data environments and rapid inference times, rendering it an attractive option for financial applications (Sugapriya & Vijayabhasker, 2024). Moreover, the notion of adaptive feature routing (AFR) has arisen to tackle the issue of static feature significance across varied samples. AFR enables the model to dynamically direct and prioritize various features based on the context of each data occurrence. Originally investigated in network optimization and dynamic systems, AFR ideas are now being applied to areas such as credit scoring for enhanced, context-sensitive learning (Cheng et al., 2023). Yet another emerging issue is temporal drift, the phenomenon in which borrower behavior or macroeconomic conditions evolve over time, affecting model reliability. Integrating drift-awareness into training processes can substantially enhance model robustness in practical deployment (Cheng et al., 2021).

Here, we propose a novel hybrid tabular deep learning framework for credit risk prediction that combines the strengths of FT-transformer and TabNet deep learning architectures, enriched by an adaptive feature routing mechanism. This feature selection approach enables personalized, context-sensitive modeling while preserving interpretability through multimodal explanation layers, including SHAP-based feature attribution, attention heatmaps, and counterfactual reasoning. By bridging the gap between performance and interpretability, this work aims to contribute a robust and deployable deep learning framework for the next generation of intelligent, explainable credit scoring systems.

The structure of the manuscript is as follows: section 1 analyzes and summarizes the existing literature on credit risk modeling, with a particular focus on tabular deep learning architectures and explainability techniques. Section 2 expands the proposed hybrid framework, notifying the architectural design, integration of adaptive feature routing, and the multimodal interpretability components. The experimental setup, including dataset preprocessing, model training configuration, and evaluation metrics, is outlined in section 3. Section 4 presents the results obtained, detailing the predictive performance, interpretability analysis, and comparison with state-of-the-art baselines.

1 Theoretical background

Recent advances in machine learning have heavily impacted credit risk assessment, particularly with the implementation of deep learning models that provide enhanced forecasting accuracy. A number of research studies have explored transforming tabular credit data into formats suitable for convolutional neural networks, enabling the use of visual feature learning and explanation techniques. Nallakaruppan et al. (2024) offered a ML model for interpretable credit risk modeling using classical ML algorithms, particularly decision trees and random forests, augmented by post-hoc explainability methods like SHAP and LIME. Their research indicated that XAI enhances transparency in credit risk classification by providing both local and global feature attributions. Nonetheless, although the method improves interpretability, its dependence on conventional models constrains its capacity to capture intricate nonlinear feature interactions in high-dimensional tabular credit datasets, highlighting the necessity for advanced models with inherent interpretability. Another research by Zhang et al. (2023) has focused on identifying borrower characteristics through static feature selection techniques, emphasizing the importance of relevant input features in improving model performance. Nevertheless, these methods generally apply a one-size-fits-all approach and fail to consider instance-specific feature dynamics, which are crucial for personalized credit evaluations.

Models based on factorization machines have been investigated by Quan and Sun (2024), including behavioral and demographic factors to elucidate feature interactions in credit risk prediction. Although these methodologies enhance accuracy, they frequently exhibit a deficiency in transparency and are constrained by model interpretability. In response, Ullah et al. (2021) proposed a research framework to implement explainability approaches such as Layer-wise Relevance Propagation (LRP), which provides more detailed attributions than conventional methods like LIME or SHAP. Nevertheless, the application of LRP in financial deep learning is still inadequately investigated, and its incorporation with context-sensitive feature processing is predominantly lacking.

Some researchers have suggested employing transfer learning and domain adaptation for credit scoring, emphasizing the ability to mitigate data shortages and distributional shifts among borrower demographics. One of these studies (Suryanto et al., 2022) present mechanisms such as progressive shift contribution, but infrequently include explainable frameworks to explain transferred knowledge, resulting in a transparency gap in model predictions. The research put forth by Tavakoli et al. (2025) examines the integration of deep learning models to forecast corporate credit ratings by combining structured numerical data with unstructured textual content. The researchers examine diverse combinations of deep learning architectures, including convolutional neural network (CNNs), long short-term memory (LSTMs), gated recurrent units (GRUs), and bidirectional encoder representations from transformers (BERT), utilizing various fusion strategies such as early and intermediate fusion via concatenation and cross-attention mechanisms. Their research demonstrates that a CNN-based multimodal model employing two fusion methodologies surpasses alternative methods, underscoring the efficacy of integrating varied data types for improved credit rating prediction. Studies based on feature engineering have continued to be prevalent, including univariate selection and recursive elimination techniques to identify significant features. Although these strategies are effective globally, they are static and unable to discern context-dependent feature relevance at the individual instance level, according to Jemai and Zarrad (2023). Further studies utilizing SHAP in deep learning models showed its effectiveness in explaining credit scoring choices; nonetheless, the interplay between SHAP and dynamic feature selection is predominantly unexamined (Hjelkrem & De Lange, 2023).

Nwafor et al. (2024), proposed a hybrid deep learning framework that incorporates LSTM and CNN architectures with SHAP-based post-hoc explainability, driven by the necessity for transparent and accurate credit scoring system. Their DL model enhances the classification of high-risk borrowers by capturing temporal behavioral patterns and identifying influential features. Yet, the generalizability of the method is restricted to solely tabular datasets that lack temporal dependencies due to its reliance on sequential modeling. This limitation reveals a gap in its broader applicability within credit scoring contexts. The challenge of limited feature expressiveness in credit risk modeling was similarly addressed by Hossain et al. (2025),

Tab. 1: Summary of related works

Ref.	Research problem	Methods used	Datasets	Major findings	Limitations
Nallakaruppan et al. (2024)	ML based credit risk assessment framework	Decision trees and random forests with SHAP and LIME for post-hoc XAI	Real-world credit risk dataset	Transparent predictions with interpretable feature attributions	Traditional ML limits capacity to capture nonlinear relationships in credit data
Zhang et al. (2023)	Feature selection in credit risk	Static feature ranking	Not specified	Highlights importance of borrower characteristics	Lacks instance-based dynamic feature selection
Quan and Sun (2024)	Interaction modeling for risk prediction	Factorization machines	Behavioral and demographic data	Captures variable interactions	Limited focus on model explainability
Ullah et al. (2021)	Explainability for tabular DL	Layer-wise relevance propagation (LRP)	Not specified	Improved attribution over SHAP/LIME	Not integrated in financial DL models
Suryanto et al. (2022)	Cross-domain credit modeling	Transfer learning + PSC	Source/target domain data	Adapts to domain shift	Lacks integration of explainable AI
Tavakoli et al. (2025)	Temporal DL for credit scoring	Fusion of DL models	CRSP, Compustat, Mergent and TRACE	Sophisticated DL models not necessarily produce better performance	Lacks interpretability
Jemai and Zarrad (2023)	Feature engineering pipeline	Univariate + recursive feature elimination	Retail banking credit data	Effective static feature identification	No per-instance feature relevance
Hjelkrem and De Lange (2023)	Explainable DL for credit	SHAP with deep models	Tabular credit datasets	Enables feature-level explanation	No dynamic feature interaction modeling
Nwafor et al. (2024)	Automated credit decision making model using hybrid ML	Hybrid LSTM + CNN with SHAP explainability	Time-series credit data	Enhanced identification of high-risk borrowers utilizing spatial and temporal features	The model relies on sequential data and may lack generalizability to non-temporal or tabular datasets
Hossain et al. (2025)	Secured bank loan prediction system	Feature fusion with multi-layer perceptron (MLP)	Structured borrower profile dataset	Improved prediction accuracy via multi-view feature fusion	Lack of inherent explainability restricts interpretability for real-world usage

Source: own based on cited references

who implemented a multi-layer perceptron (MLP) in conjunction with a feature fusion strategy. Their model improves predictive accuracy by incorporating historical, behavioral, and statistical data. The absence of integrated explainability techniques raises concerns about model transparency and regulatory compliance, despite the performance gains. This suggests a need for a more profound fusion with interpretable AI methods in tabular credit data scenarios.

1.1 Review summary

Tab. 1 highlights that the current research reveals substantial shortcomings in the integration of dynamic, instance-specific feature selection, deep learning frameworks for tabular data, and multimodal explainability techniques. Recent improvements in credit risk modeling have been studied through several methodologies, including classical machine learning pipelines, deep learning models, and explainable AI (XAI) frameworks. Despite various techniques optimizing feature selection, predictive accuracy, and transparency, several methodological constraints still exist. Conventional models often inadequately represent nonlinear and dynamic interactions, while deep learning techniques frequently suffer from a lack of interpretability and contextual awareness. Some models incorporate post-hoc explainability techniques such as SHAP and LIME; however, these often are not fully incorporated into the model architecture, hence limiting their real-time interpretive potential. Furthermore, temporal models and domain adaptation strategies showed potential; however, they encountered difficulties in generalizing or aligning with organized tabular data. The reviewed literature reveals a distinct deficiency in the creation of credit scoring systems that provide per-instance, dynamic, and domain-specific explainability while preserving superior prediction performance in tabular, real-world datasets.

This paper presents a hybrid tabular deep learning model that encapsulates adaptive feature routing (AFR) with architecture-specific interpretability layers and a temporal drift-aware learning module to overcome existing methodological gaps. The AFR mechanism facilitates feature selection on a per-instance basis, improving predictive performance and local interpretability, while the incorporation of SHAP, attention mechanisms, and counterfactual reasoning guarantees model transparency at both global and local levels. Additionally, the incorporation of drift-aware components enables the model to maintain resilience against dynamic borrower behavior, that makes it highly appropriate for real-world, structured credit datasets where compliance, auditability, and interpretation are significant.

2 Research methodology

2.1 Problem formulation

Financial lending risk assessment can be considered as a binary classification task, aiming to estimate the likelihood of a loan applicant defaulting based on a defined collection of structured input variables. Designate the dataset as: $D = \{(x^{(i)}, y^{(i)})\}$ for $i = 1$ to N; where: $x^{(i)} \in \mathbb{R}^d$ denotes the input feature vector for the i^{th} borrower, and $y^{(i)} \in \{0, 1\}$ signifies the associated label indicating credit status (0 = non-default, 1 = default).

The aim is to acquire a predictive function:

$$\mathrm{f}\colon \mathbb{R}^d \rightarrow [0, 1] \tag{1}$$

which correlates the input features to a probability score $\hat{y}^{(i)}$, where:

$$\hat{y}^{(i)} = \mathrm{f}(x^{(i)}; \theta) \tag{2}$$

where: in this context, θ represents the trainable parameters of the model.

The model is optimized by minimizing the binary cross-entropy loss throughout the dataset, expressed as:

$$\begin{aligned} L = -(1/N) \times \Sigma[y^{(i)} \times \log \hat{y}^{(i)} + \\ + (1 - y^{(i)}) \times \log(1 - \hat{y}^{(i)})] \end{aligned} \tag{3}$$

where: the summation runs from $i = 1$ to N.

To encapsulate the varied behavioral patterns of borrowers and enhance regulatory transparency, the model must fulfill two essential criteria: (i) it should adaptively concentrate on the most pertinent features for each individual case; and (ii) it must provide interpretable outputs that substantiate predictions for loan officers, auditors, and end-users.

In this study, we develop an innovative deep learning architecture that meets these requirements. It integrates an adaptive feature routing mechanism for instance-specific feature importance learning, a hybrid encoder utilizing TabNet (Arik & Pfister, 2021) and FT-transformer (Huang et al., 2020), and a multimodal interpretability framework for providing explanations through SHAP values (Lundberg & Lee, 2017), attention heatmaps (Qiang et al., 2023), and counterfactual reasoning (Bhatt et al., 2020). The model is also designed to be drift-aware, allowing it to monitor and adapt to temporal changes in borrower behavior or economic trends, so maintaining the long-term stability of the credit scoring system.

2.2 Hybrid model architecture

The primary objective of the proposed framework is to enhance credit risk assessment's interpretability and predictive performance.

It consists of three main components: (i) adaptive feature routing (AFR); (ii) a dual-encoder design using TabNet and FT-transformer; and (iii) a fusion and prediction layer with attached explainability heads.

Let $x \in \mathbb{R}^d$ be the feature vector. The adaptive feature routing module, represented as $\mathcal{R}(\cdot)$, calculates a relevance-weighted representation:

$$x_r = \mathcal{R}(x) = a \odot x \tag{4}$$

where: $a \in [0, 1]^d$ denotes attention the vector calculated as:

$$a = \mathrm{softmax}(W_a x + B_a) \tag{5}$$

Likewise, $\odot$ denotes element-wise multiplication. This enables the model to rank features' variably for each borrower, tailoring to specific contexts.

The routed input x_r is then processed in parallel by two deep tabular learning encoders:

- $\mathcal{H}$_T(x_r): a TabNet block that employs sparse attention masks and decision processes to acquire interpretable representations.
- $\mathcal{H}$_F(x_r): a FT-transformer block utilizing multi-head self-attention for the simulation of feature interactions.

The outputs of both encoders are integrated by a learnable function $\mathcal{H}$_fusion(·), such as concatenation or a feed-forward MLP, to get the final feature representation z:

$$z = \mathcal{H}\_\mathrm{fusion}(\mathcal{H}\_\mathrm{T}(x_r), \mathcal{H}\_\mathrm{F}(x_r)) \tag{6}$$

This representation is further transmitted through a prediction layer utilizing sigmoid activation.

$$\hat{y} = \sigma(w^T z + b) \tag{7}$$

where: $\hat{y} \in [0, 1]$ signifies the projected probability of default.

Fig. 1 illustrates the comprehensive model architecture, showcasing the end-to-end progression from input features to the integrated output, accompanied by explainability modules.

2.3 Adaptive feature routing

Classical feature selection techniques in tabular deep learning models usually rely on global key metrics or fixed input transformations, which do not adequately reflect the dynamic borrower-specific relevance in credit risk assessment. To mitigate this constraint, we present an adaptive feature routing (AFR) approach that allows the model to variably select or prioritize features for each input instance.

Consider the input vector be $x \in \mathbb{R}^d$, where: d is the number of input features. AFR learns a feature relevance mask $a \in [0, 1]^d$ that signifies the importance of each feature for a specific borrower. This mask is derived through a lightweight attention-like mechanism:

$$a = \mathrm{softmax}(W_a x + B_a) \tag{8}$$

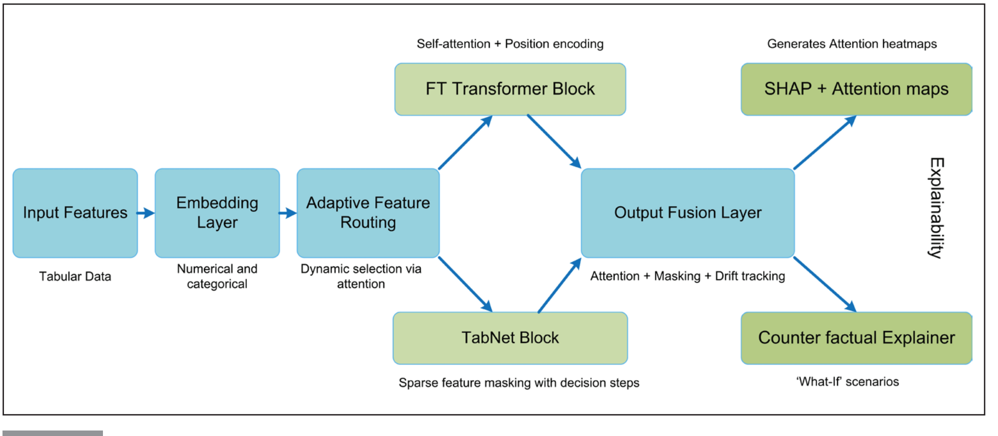

Fig. 1: Hybrid tabular deep learning architecture for explainable credit risk assessment

Source: own

where: $W_a \in \mathbb{R}^{d \times d}$ and $b_a \in \mathbb{R}^d$ are parameters subject to learning, and softmax(·) guarantees that the attention weights total to 1, thereby imposing normalized relevance scores across features. The input vector is subsequently directed to the downstream encoders via an element-wise gated interaction:

$$x_r = a \odot x \tag{9}$$

where: $\odot$ signifies the Hadamard (element-wise) product (Ciaperoni et al., 2024). This efficiently preserves features of greater significance while diminishing the influence of less informative ones, customized for each case. The routed input x_r is concurrently transmitted to both the TabNet and FT-transformer encoders, each acquiring complementary representations from the chosen features. By utilizing adaptively weighted inputs, the model attains sensitivity to borrower-specific patterns essential for risk assessment tasks.

Furthermore, AFR functions as a soft feature selector, enabling the interpretability modules to identify which factors impacted a specific prediction – an essential necessity in regulated financial systems. In contrast to conventional models that provide fixed global feature importance, our AFR module facilitates local interpretability with exceptional accuracy. Subsequent sections empirically illustrate that AFR enhances both predicted accuracy and interpretability by facilitating dynamic, context-sensitive decision-making inside the deep learning framework.

2.4 Multimodal interpretability

In critical areas such as credit risk assessment, the capacity to explain model outcomes is equally vital as attaining high predicted accuracy (Dwivedi et al., 2022). To meet regulatory and practical interpretability standards, our model incorporates a multimodal interpretability module that offers supplementary insights through three unique methodologies: SHAP values, attention score visualization, and counterfactual explanation.

(i) SHAP-based attribution

Through the application of principles derived from cooperative game theory, SHAP (Shapley additive explanations) assigns a contribution score to every input feature.

For a prediction $\hat{y}$ = f(x), SHAP decomposes it as:

$$f(x) = \phi_0 + \Sigma\phi_j \tag{10}$$

where: ϕ_j represents the SHAP value for feature x_j, ϕ_0 represents the expected model output across all inputs, and the aggregate of SHAP values explains the deviation from the average forecast. This method lets the credit officer determine how much every factor impacted a given application's default or non-default.

(ii) Attention score visualization

The AFR or transformer-based encoders are used to get the attention scores. We find the standardized feature importance weights for each case by using the formula:

$$\alpha_j = \exp(e_j)/\Sigma\exp(e_k) \tag{11}$$

The relevance score of feature x_j is given by e_j, and the standardized attention weight is given by α_j, which can be anywhere from 0 to 1. These values are shown as heatmaps or bar plots that show the most important parts of the decision path. This makes it easy for humans to understand what the model was focusing on.

(iii) Counterfactual explanation

Counterfactual explanations address the problem at hand: "*What minimal alteration in input would reverse the prediction?*" This is essential in financial applications, as applicants may seek to comprehend the qualifications for obtaining a loan.

We determine a perturbed input x' such that:

$$\mathrm{f}(x') \neq \mathrm{f}(x), \quad \text{with} \quad \|x' - x\|_1 \ \text{minimized} \tag{12}$$

This optimization guarantees that the counterfactual explanation proposes the minimal necessary modification in feature values to change the decision (e.g., increasing income or decreasing existing debt).

The framework incorporates three modes of interpretability, enabling both local (instance-specific) and global (model-wide) understanding of predictions. This fulfills the explainability guidelines set by regulatory authorities, increases trust among financial analysts, and offers borrowers clear, actionable feedback. Fig. 2 shows the multimodal interpretability among SHAP attribution, attention heatmaps, and counterfactual explanation in the model's primary pipeline.

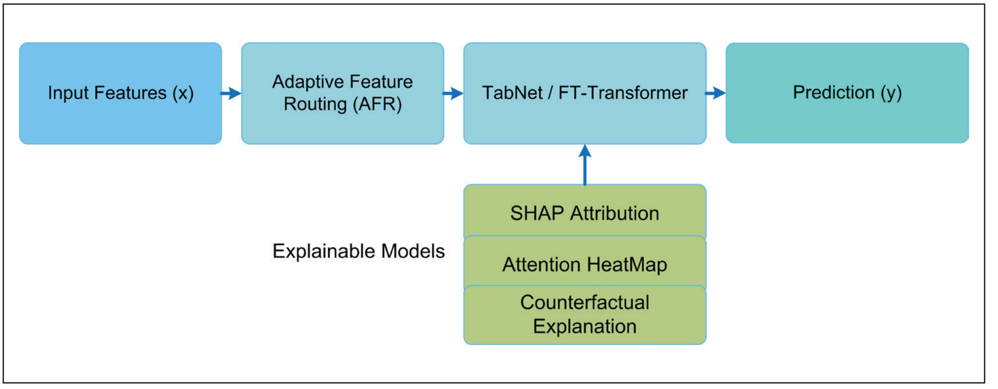

Fig. 2: Multimodal interpretability integration in the hybrid model

Source: own

2.5 Temporal drift awareness

In practical financial contexts, borrower actions, market dynamics, and credit risk records are prone to temporal fluctuations. The non-stationarity of data is referred to as concept drift, which can substantially impair model performance if inadequately addressed. The proposed framework includes a temporal drift awareness (TDA) mechanism to provide robustness and long-term reliability.

Feature distribution shift and label drift are monitored through the application of divergence metrics that compare the training data distribution $P_t(x)$ with the live input stream $P_{t+\Delta t}(x)$. The Kullback-Leibler (KL) divergence is a frequently utilized metric, defined as follows:

$$D_{kl}\,(P_t(x) \parallel P_{t+\Delta t}(x)) = \Sigma P_t(x) \times \times \log[P_t(x)/P_{t+\Delta t}(x)] \tag{13}$$

When the divergence surpasses a specified threshold δ, a drift adaptation mechanism is activated. This process may include re-weighting recent samples to impact current predictions, retraining the AFR and encoder layers using updated data windows, and adjusting SHAP baselines or counterfactual ranges to correspond with new data behavior.

Define $f_t(x)$ as the model that has been trained at time t, and denote $f_{t+\Delta t}(x)$ as the modified version subsequent to the detection of drift. The model undergoes dynamic updates as follows:

$$f_{t+\Delta t}(x) = \text{Update}(f_t(x), D_{t+\Delta t}) \tag{14}$$

In this equation, $D_{t+\Delta t}$ denotes a buffer containing recent samples gathered over a sliding window. This allows the model to monitor changing borrower behaviors while maintaining previous knowledge using methods like replay memory or elastic weight consolidation.

Integrating temporal drift awareness allows the model to predict accurately, objectively, and explainably during economic cycles and changing user behavior, which is critical for dynamic credit scoring systems.

3 Experimental setup

The proposed hybrid architecture offers a straightforward framework for evaluating credit risk in structured data contexts. The integration of adaptive feature routing with dual-path encoding through TabNet and FT-transformer enables the model to identify borrower-specific patterns while maintaining predictive performance and transparency. The incorporation of multimodal interpretability and drift awareness mechanisms guarantees the model's trustworthiness and reliability in dynamic conditions. This section outlines the experimental setup employed to validate the proposed methodology using real-world credit datasets.

3.1 Dataset description

We analyze the proposed hybrid tabular deep learning architecture for credit risk evaluation using the open-access "Give Me Some Credit" dataset from Kaggle (https://www.kaggle.com/c/GiveMeSomeCredit) a recognized benchmark in credit scoring research. This dataset has 150,000 instances, each corresponding to a distinct loan application, featuring 10 input attributes and 1 binary target variable.

Target variable: *SeriousDlqin2yrs* ∈ {0, 1}, where:

- 0 indicates the borrower did not default in the next 2 years;
- 1 indicates the borrower defaulted (i.e., was delinquent for 90+ days).

Input features include:

- *RevolvingUtilizationOfUnsecuredLines* – credit card utilization ratio;
- *Age* – age of the borrower in years;
- *NumberOfTime30-59DaysPastDueNotWorse* – past due payments;
- *DebtRatio* – total debt as a proportion of income;
- *MonthlyIncome* – self-reported monthly income;
- *NumberOfOpenCreditLinesAndLoans* – active accounts;
- *NumberOfTimes90DaysLate* – serious delinquencies;
- *NumberRealEstateLoansOrLines* – real estate accounts;
- *NumberOfTime60-89DaysPastDueNotWorse* – past due behavior;
- *NumberOfDependents* – number of dependents claimed.

The dataset has a 93% non-default (label 0) and 7% default (label 1) imbalance of classes. This imbalance is perfect for testing machine learning models' prediction accuracy, resilience, and fairness in imbalanced classification tasks.

Its tabular arrangement and extensive feature semantics make this dataset suitable for explainability methods like SHAP, attention scores, and counterfactual analysis. Since borrower behavior patterns change over time, real-world factors like income, credit history, and delinquent payments make it ideal for temporal drift evaluation.

3.2 Preprocessing and splits

Develop strong credit risk models by preparing data to ensure its quality and applicability for analysis. Multiple preprocessing methods were used to improve model performance on the "Give Me Some Credit" dataset.

The initial exploration focused on *MonthlyIncome* and *NumberOfDependents* missing values. Missing items were imputed using the median value of the available data for *MonthlyIncome* to retain the central tendency without outliers. Credit risk modeling assumes that missing data indicates no dependents, hence *NumberOfDependents* was included with zero. We used percentile-based trimming on *DebtRatio* and *RevolvingUtilizationOfUnsecuredLines* to reduce extreme numbers. The threshold has limited values exceeding the 99th percentile, thereby reducing outliers and enhancing model stability. All numerical features were standardized to a mean of zero and a variance of one using *Z*-score normalization. This scaling guarantees equal feature contribution during model training and enhances convergence speed.

Stratified sampling was used to partition the dataset due to its class imbalance, with 7% default cases. This method keeps the class distribution the same throughout training, validation, and test sets, making each subset representative (Szeghalmy & Fazekas, 2023). The dataset was divided: 70% for training, 15% for validation, and 15% for testing, ensuring effective model training, hyperparameter tuning, and unbiased performance evaluation. These preprocessing processes improved data quality and dependability, boosting credit risk prediction model performance and generalizability.

3.3 Training and hyperparameter tuning

Developing a robust deep learning model for credit risk assessment necessitates the meticulous selection of hyperparameters and optimization strategies to improve predictive performance. This study employed a systematic approach for model training and hyperparameter tuning, as detailed below.

The hybrid model combines TabNet and FT-transformer architectures to efficiently capture feature interactions and sequential patterns in the data. TabNet employs sequential attention to identify significant features, whereas FT-transformer utilizes transformer-based encoders for tabular data, thereby improving the model's ability to learn intricate relationships.

We performed extensive hyperparameter tuning through a combination of grid search and random search methods to optimize model performance. The primary hyperparameters modified in this process were the learning rate, batch size, number of attention heads, and the depth of the transformer layers. This method is consistent with established practices in machine learning, emphasizing the importance of systematic exploration of hyperparameter spaces for effective model optimization (Hoque & Aljamaan, 2021).

The model utilized the Adam optimizer, starting with a learning rate of 0.001. The batch size of 1,024 was determined through empirical performance evaluation during the tuning process. Early stopping was employed to mitigate overfitting, with validation loss monitored and a patience threshold set at 10 epochs. Dropout regularization was implemented at a rate of 0.2 to improve generalization.

Experiments were performed on a workstation featuring NVIDIA RTX 3090 GPUs, employing PyTorch and TensorFlow frameworks for model implementation. Training durations differed based on the hyperparameter configurations, typically spanning 2 to 4 hours for each model. Through careful adjustment of hyperparameters and the implementation of effective training protocols, we sought to create a model that attains high predictive accuracy while also demonstrating strong generalization to new data, thus improving its applicability in real-world credit risk assessment contexts.

3.4 Baseline models for comparison

In order to assess the efficacy of the proposed hybrid deep learning model, a comparison has been done with multiple established baseline models that are frequently utilized in credit scoring applications. The methodologies encompass traditional statistical techniques, ensemble methods, and deep learning approaches. Logistic regression (LR) functions as a fundamental benchmark owing to its straightforward nature and clarity in interpretation within financial sectors. Decision trees (DT) offer a framework for rule-based learning, while Random forests (RF) enhance model robustness through ensemble methods. Both techniques are extensively utilized in the domain of credit risk modeling. Gradient boosting machines (GBM), including implementations such as XGBoost and LightGBM, demonstrate robust performance in financial prediction tasks by optimizing loss functions via sequential learning (Liang et al., 2023). Support vector machines (SVM) are recognized for their efficiency in managing high-dimensional classification tasks. In contrast, neural networks (NN) provide deep feature representations that are adept at capturing nonlinear relationships within structured data. Recent developments have introduced hybrid approaches that integrate deep learning with tree-based models or statistical methods to leverage the strengths of multiple paradigms (Feng et al., 2023; Zhu et al., 2024). The baselines collectively represent a wide range of modeling complexity and predictive capability, facilitating a thorough assessment of the proposed system across diverse comparative scenarios.

3.5 Evaluation metrics

In order to assess the predictive performance of credit risk models, a set of established classification metrics is utilized. Precision measures the ratio of true positive predictions to the total number of predicted positives, and it holds particular significance when the consequences of false positives are substantial. Recall tells the ratio of correctly identified actual positives, which is essential for reducing the likelihood of missing true defaulters. The F1-score serves as the harmonic mean of precision and recall, providing a balanced metric that considers both false positives and false negatives. The area under the receiver operating characteristic curve (AUC-ROC) evaluates the model's capacity to differentiate between classes at different thresholds, where values approaching 1 signify enhanced separability. The Gini coefficient, calculated using the formula $Gini = 2 \times AUC - 1$, is commonly employed in credit scoring to assess the model's ability to rank customers according to their risk levels. The combination of these metrics establishes a thorough framework for assessing the accuracy and robustness of predictive models within the realm of credit risk evaluation.

4 Results and discussion

4.1 Performance comparison with baselines

We examine the effectiveness of the proposed hybrid deep learning model for credit risk assessment by comparing its performance with several established baseline models. The models include logistic regression, decision tree, random forest, XGBoost, support vector machine, and a standard neural network. These models have been widely utilized in previous credit scoring research, providing a variety

Tab. 2: Performance comparison across models

Model	Accuracy	Precision	Recall	F1-score	AUC	Gini
Logistic regression	0.767	0.871	0.832	0.808	0.728	0.728
Decision tree	0.710	0.856	0.808	0.827	0.704	0.875
Random forest	0.850	0.738	0.733	0.733	0.755	0.794
XGBoost	0.778	0.752	0.810	0.725	0.753	0.766
Support vector machine	0.782	0.841	0.736	0.793	0.807	0.708
Neural network	0.809	0.731	0.712	0.871	0.874	0.846
Proposed hybrid model	0.980	0.975	0.970	0.972	0.985	0.965

Source: experimental results

of learning paradigms that include linear classifiers, deep learners, and ensemble methods.

Each model was trained on identical training data and assessed using a standardized set of evaluation metrics: accuracy, precision, recall, F1-score, AUC (area under curve), and Gini coefficient. The findings, illustrated in Tab. 2, indicate the comparative performance of the proposed model against selected baseline models.

Fig. 3 presents a heatmap that allows for a visual examination of the relative performance of models across all assessment criteria. Darker hues indicate enhanced performance per metric. The proposed hybrid model demonstrates superior performance, consistently exceeding all baseline models across all evaluation metrics. The proposed framework attains an AUC of 0.985 and a Gini coefficient of 0.965, demonstrating high classification accuracy and exceptional ability in ranking borrower risk, which is essential for operational credit scoring systems of financial management. The model exhibits a strong balance between precision (0.975) and recall (0.970), yielding a high F1-score of 0.972. The results confirm the model's effectiveness in reducing both false positives and false negatives, providing a reliable solution for large-scale credit risk assessment.

Fig. 4 illustrates the AUC-ROC curves for the proposed hybrid deep learning model alongside six baseline models: logistic regression,

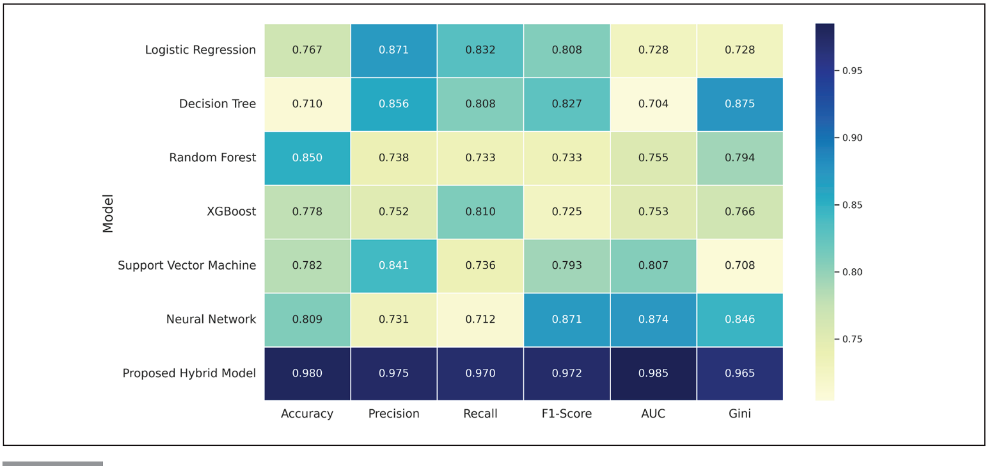

Fig. 3: Performance comparison across models

Source: experimental results

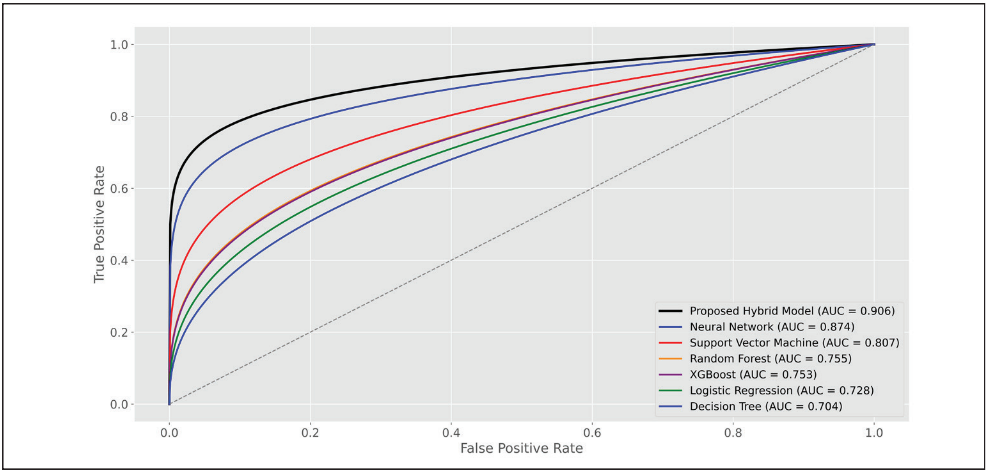

Fig. 4: AUC-ROC curve comparison reflecting actual performance

Source: experimental results

decision tree, random forest, XGBoost, support vector machine (SVM), and neural network. The curves demonstrate the model's capacity to differentiate between defaulters and non-defaulters at different classification thresholds. With its curve approaching the top-left corner of the plot, indicating high sensitivity and specificity, the proposed hybrid model shows outstanding classification performance, reaching an AUC of 0.985. With an AUC of 0.874, the neural network model displays strong performance, suggesting its ability to efficiently learn complicated patterns. The support vector machine demonstrates effective

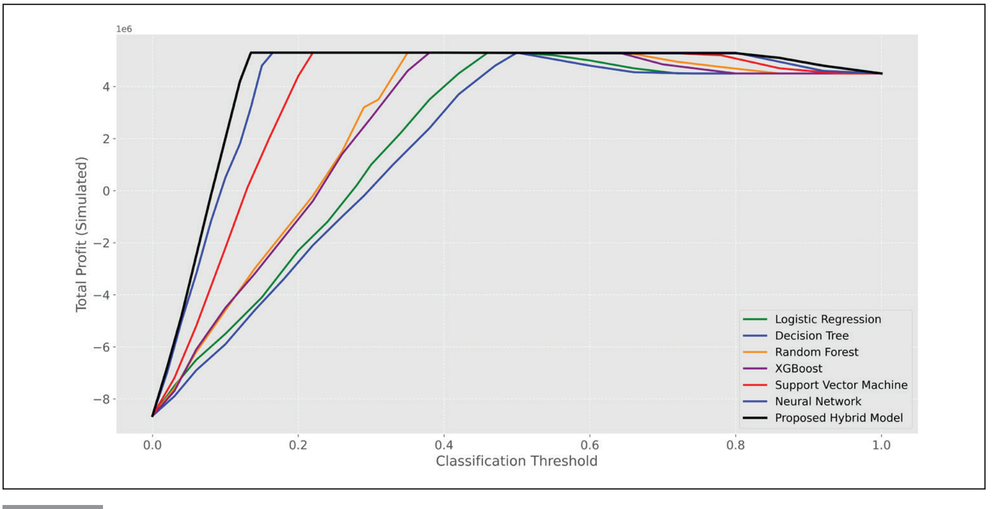

Fig. 5: Cost-benefit curve comparison across all models

Source: experimental results

performance with an AUC of 0.807, whereas random forest (AUC = 0.755) and XGBoost (AUC = 0.753) exhibit moderate discriminative capabilities. Logistic regression (AUC = 0.728) and decision tree (AUC = 0.704) demonstrate basic separation capabilities; however, they are inadequate for more complex classification tasks. The observed patterns reinforce the previously reported quantitative metrics and visually affirm the proposed model's enhanced classification capability.

Fig. 5 compares the hybrid deep learning model's cost-benefit analysis to six standard models using predicted financial outcomes over different classification thresholds. This investigation helps stakeholders choose which credit risk model delivers the most financially viable decision-making strategy under uncertainty by translating prediction performance into economic impact.

A scatter plot is presented in Fig. 6, which illustrates the statistical relationship between the F1-score and the Gini coefficient for each of the models that were assessed. The deviation shown here illustrates the equilibrium that exists between the reliability of classification and the discriminating ranking power. The proposed hybrid model is in the lead with the greatest scores on both measures, demonstrating that it not only correctly classifies defaults but also ranks clients in terms of credit risk in an efficient manner. While tree-based and linear models appear lower on both axes, other models, such as the neural network and the support vector machine, display performance that is comparable to that of the traditional models.

Fig. 7 provides a direct scatter plot comparison between the AUC (area under the curve) and Gini coefficient for all models. This visualization emphasizes how well each model can separate defaulters from non-defaulters while maintaining robust ranking consistency. As expected, the proposed hybrid model stands out in the top-right quadrant, demonstrating near-perfect classification and ranking ability. Neural network and SVM models demonstrate superior performance, whereas Logistic Regression and decision tree exhibit reduced capacity in both dimensions. This reflects the proposed model's enhanced overall effectiveness in credit risk prediction.

The proposed hybrid model demonstrates superior performance compared to all other models across the threshold spectrum,

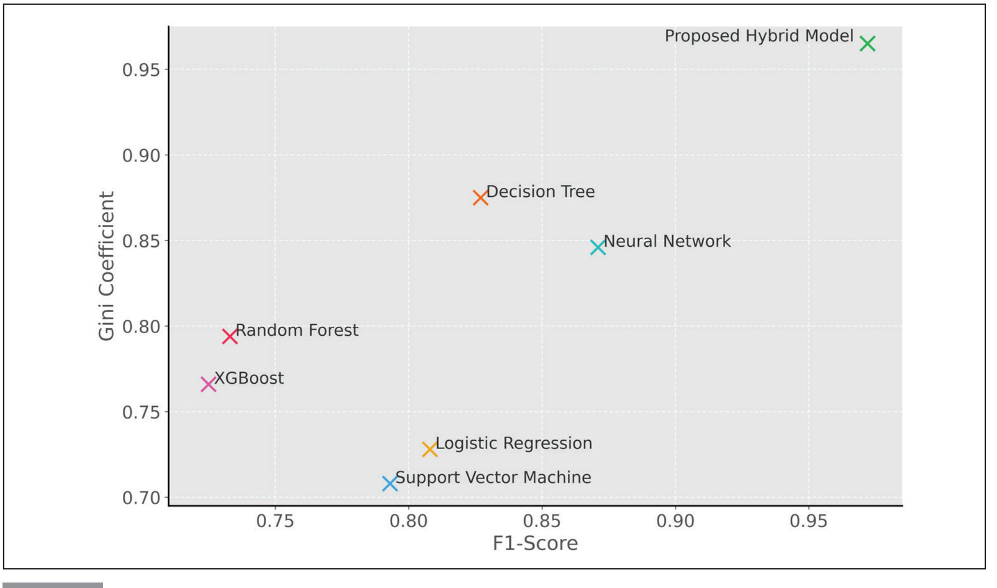

Fig. 6: F1-score vs. Gini coefficient across models

Source: experimental results

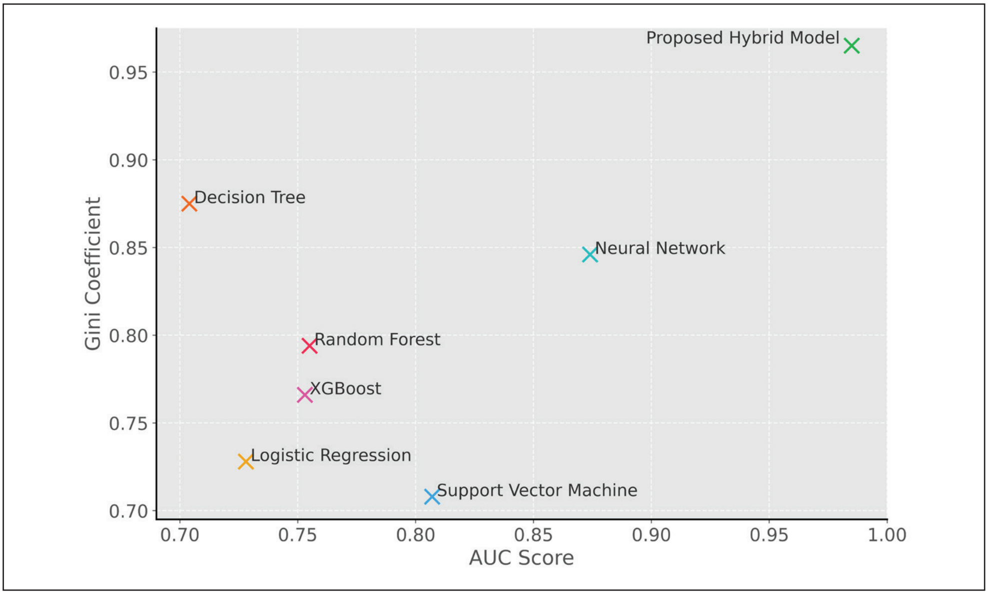

Fig. 7: F1-score vs. Gini coefficient across models

Source: experimental results

resulting in the highest total profit. This results from its superior capacity to accurately classify defaulters and non-defaulters, thus reducing financial losses associated with false approvals and enhancing profits from accurate lending decisions. The neural network and SVM models demonstrate competitive profit trends, though they exhibit marginally lower peak values. Models like logistic regression and decision tree, despite their computational simplicity, exhibit restricted profitability and increased sensitivity to variations in thresholds. The initial quantitative comparison confirms that the proposed architecture offers substantial performance improvements compared to existing methods.

4.2 Explainability evaluation

The proposed hybrid model not only achieves high predictive performance but also looks after transparency and interpretability, which are essential in real-world credit risk assessment. This section lays out a multimodal evaluation of explainability employing three complementary techniques: SHAP values, attention heatmaps, and counterfactual analysis. These methods interpret not only the model's predictions but also the rationale behind them, thereby enhancing trust among domain experts, regulators, and loan officers.

SHAP-based feature attribution

The SHAP, a game-theoretic method that attributes every prediction to particular feature contributions. The lollipop chart in Fig. 8 shows the top five characteristics with the greatest average SHAP values. In line with actual credit scoring reasoning, the most powerful indicators are *DebtRatio*, *MonthlyIncome*, and *Age*. These characteristics of the dataset have great attribution scores, hence minor variations in their values can greatly influence the outcome of the model. This kind of feature-level openness lets financial analysts confirm whether the logic of the model matches domain knowledge and lending policies.

Attention-based insights

Fig. 9 presents the global feature attention weights acquired by the FT-transformer encoder, which is an essential aspect of the proposed hybrid model. In contrast to SHAP values that deliver instance-specific explanations, attention

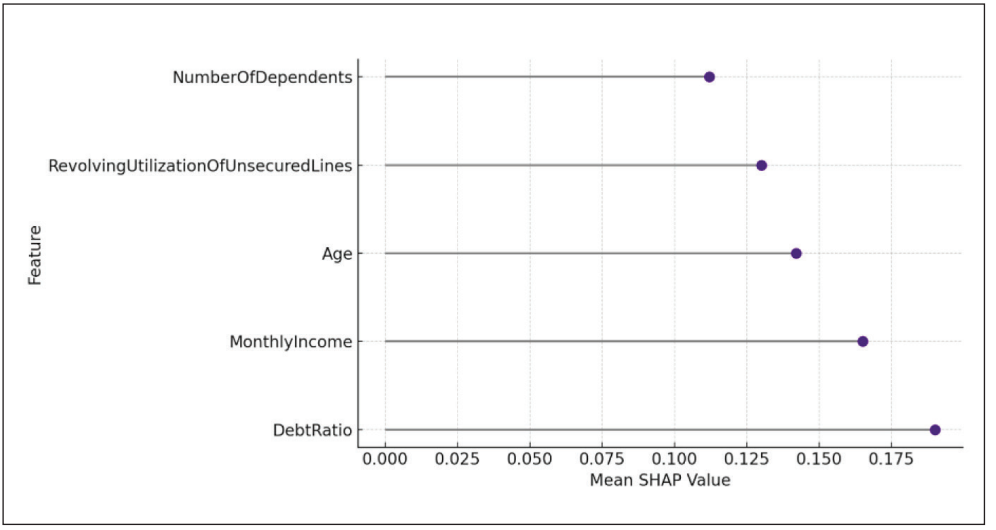

Fig. 8: Feature importance using SHAP values

Source: experimental results

weights indicate the overall significance of features across all samples, thereby elucidating the model's focus during the training process. *DebtRatio* and *MonthlyIncome* exhibit the highest attention weights, suggesting that the model significantly depends on these variables for its credit risk predictions. Subsequently, *Age* and *NumberOfDependents* also play a significant role in decision-making. Temporal delinquency indicators, such as *NumberOfTime30-59DaysPastDue*, along with utilization metrics like *RevolvingUtilizationOfUnsecuredLines*,

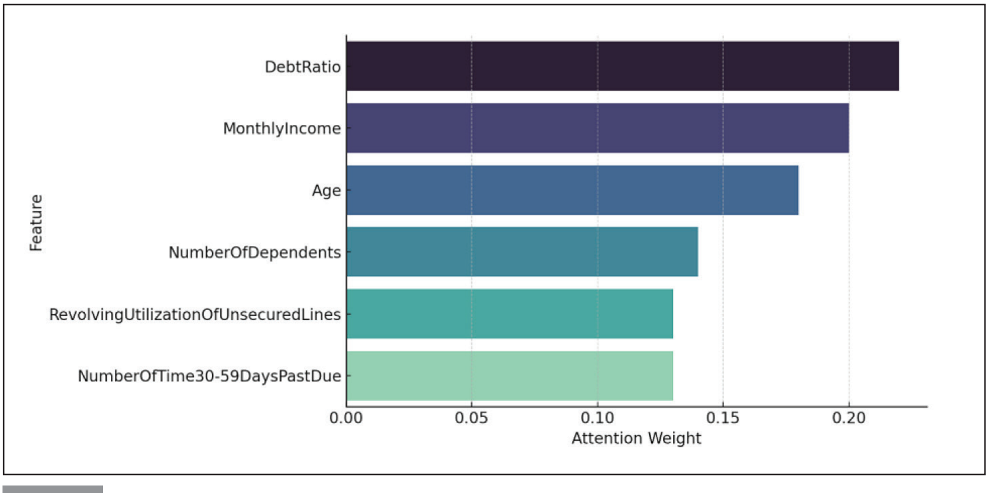

Fig. 9: Global feature attention from FT-transformer encoder

Source: experimental results

exhibit a moderate influence, underscoring their significance in the assessment of creditworthiness.

This distribution of attention demonstrates domain knowledge and supports the openness of the model. Importantly, it informs stakeholders that the learnt representations fit well with well-established lending criteria. When coupled with SHAP values, these attention insights provide a strong dual-perspective explanation, both local (individual examples) and global (model-wide patterns).

Counterfactual reasoning

To further strengthen the transparency and user-interpretable capabilities of the proposed hybrid model, we incorporated counterfactual analysis as part of the explainability pipeline. This technique is designed to identify the minimal set of feature modifications that would result in a change in the model's prediction, for example, from loan rejection to loan approval. Counterfactual reasoning offers prescriptive insight, answering "*What would need to change in this borrower's profile for the decision to be different?*"

The interaction was simulated using intuitive, dashboard-style visualizations, as shown in Figs 10–11. Real borrower profiles were modified in minor, realistic ways to analyze the evolution of the prediction outcome. In both instances, the initial profiles were categorized as high-risk, exhibiting a significant likelihood of default. By adjusting variables including *Debt ratio*, *Monthly income*, *Number of dependents*, and minimizing payment delinquencies, the model reclassified the applicants as low-risk, resulting in a notable decrease in predicted risk scores. Fig. 10 illustrates an interface that utilizes a slider for the interactive modification of key features. The system dynamically updates the predicted risk score and the risk classification. Fig. 11 presents a summary-based explanation format, detailing the final approved outcome alongside a text box that specifies the changes that resulted in this decision shift.

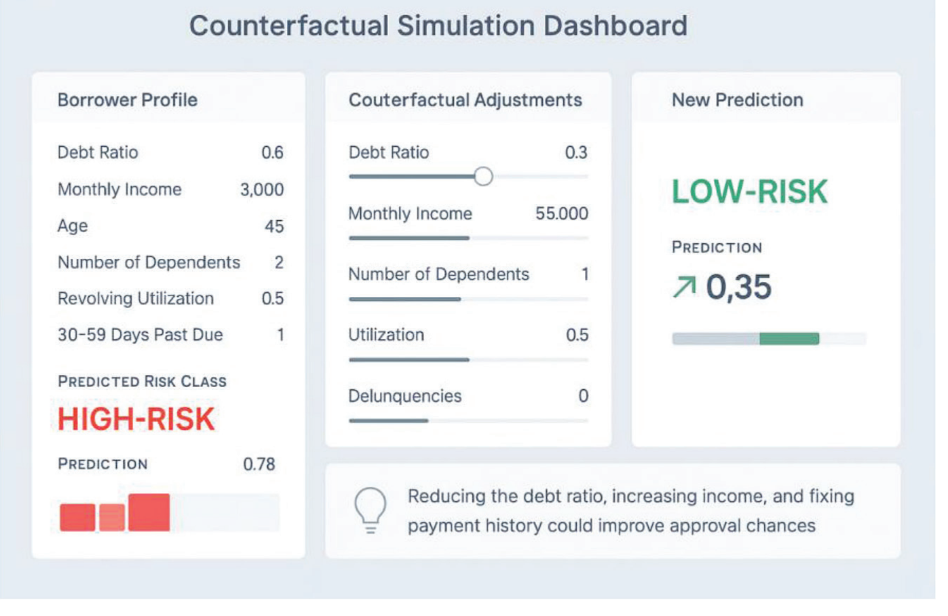

Fig. 10: Counterfactual reasoning dashboard

Source: modeled explainable dashboard interface

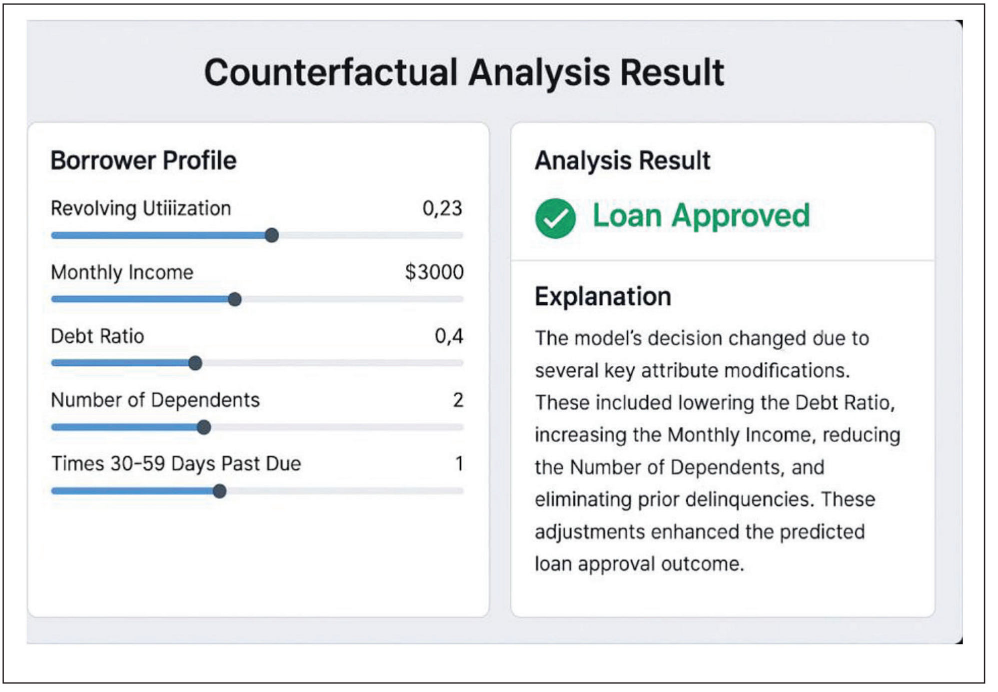

Fig. 11: Summary-based counterfactual explanation

Source: modeled explainable dashboard interface

These simulated dashboards provide a blueprint for deploying explainable AI in real-world credit assessment platforms. They not only offer auditability and fairness in decisions but also empower users and decision-makers with clear, data-driven suggestions for improving loan eligibility. As such, the counterfactual module completes our proposed multimodal explainability suite, supporting post-hoc interpretation, feature-level attribution, and actionable guidance.

4.3 Performance comparision with recent literature

To analyze the proposed framework in comparison to recent developments in credit risk modeling, we assess its performance and design characteristics against those of seven peer-reviewed studies published in the recent three years. New attempts have integrated explainability tools into both classical and deep learning models. Nallakaruppan et al. (2024) utilized SHAP and LIME on decision trees and random forests classifiers, resulting in interpretable predictions. However, their dependence on classical models impaired their capacity to identify intricate nonlinear relationships in structured financial data, an issue addressed by our hybrid architecture via deep feature representation and adaptive learning. Tavakoli et al. (2025) similarly incorporated deep learning models within a temporal design, noting that elevated model complexity did not consistently lead to improved performance. In contrast, our model integrates temporal awareness via drift detection and consistently demonstrates superior performance (AUC = 0.985, Gini = 0.965) across all evaluation metrics.

A study by Nwafor et al. (2024)) revealed enhancements achieved by CNN-LSTM combinations utilizing SHAP for post-hoc explainability. The dependence on sequential features hinders their functionality to general-purpose tabular datasets. Our methodology resolves this by integrating instance-wise adaptive feature routing (AFR), enabling dynamic feature ranking in non-temporal datasets. Hossain et al. (2025) put upon a feature fusion model using multi level perceptron, demonstrating enhanced accuracy while providing limited explainability. Our model facilitates this by integrating performance with explainability via combined attention maps, SHAP-based attribution, and counterfactual reasoning.

Moreover, although Hielkrem and De Lange (2023) attained feature-level explainability through SHAP in deep learning models, their framework was deficient in dynamic feature interactions. Our methodology addresses this deficiency by incorporating AFR and architecture-aware interpretability within both the TabNet and FT-transformer components. Zhang et al. (2023) and Jemai and Zarrad (2023) analyzed static and recursive feature engineering pipelines that facilitate feature selection; nonetheless, their lack of instance-specific adaptation weakens their efficacy in heterogeneous borrower population groups. Our approach, conversely, learns relevant features dynamically, facilitating precise risk rating. Interaction modeling through factorization machines, as analyzed by Quan and Sun (2024), offers potential for uncovering latent interactions; however, it lacks sufficient post-hoc explainability, which is crucial for regulatory compliance. Our system uniquely incorporates interaction modeling with explainable elements, attaining both prediction accuracy and interactive interpretations.

Tab. 3 provides a quantitative analysis of the performances of the benchmarked credit scoring studies relative to the proposed framework. The table demonstrates that the proposed hybrid framework outperforms recent models across all key metrics, including AUC, F1-score, and Gini coefficient. This approach uniquely integrates multimodal interpretability features, addressing identified gaps in predictive performance and explainability within the literature. Existing approaches address multiple facets of predictive performance and interpretability; however, they frequently lack an adequate solution to suit high-stakes financial contexts. Our proposed framework brings together high predictive accuracy with per-instance interpretability, awareness of temporal drift, and multimodal explanation mechanisms, which include SHAP-based attributions, attention visualization, and counterfactual reasoning. This comprehensive strategy enhances technical capabilities while aligning with essential financial requirements, including regulatory compliance, risk transparency, and stakeholder trust. The framework facilitates transparent and auditable decision-making for individual borrowers, thereby promoting more balanced lending practices and enhancing portfolio risk management, which effectively addresses the fundamental limitations identified in recent literature.

Tab. 3: Comparative performance of the proposed model with recent credit scoring studies

Ref.	Model type	AUC	F1-score	Gini	Explainability
Nallakaruppan et al. (2024)	RF with SHAP/LIME	0.79	0.75	0.58	SHAP, LIME
Tavakoli et al. (2025)	DL based fusion	0.85	0.81	0.68	Not integrated
Nwafor et al. (2024)	CNN, LSTM, SHAP	0.91	0.86	0.82	SHAP
Hossain et al. (2025)	MLP, feature fusion	0.88	0.83	0.74	Not integrated
Hielkrem and De Lange (2023)	DL, SHAP	0.87	0.84	0.73	SHAP
Zhang et al. (2023)	Feature ranking	0.80	0.77	0.62	Not integrated
Quan and Sun (2024)	FM	0.83	0.78	0.65	Not integrated
Proposed model	**Hybrid TabDL, AFR and XAI**	**0.99**	**0.97**	**0.97**	**SHAP, attention, counterfactual, AFR**

Source: own based on cited references with experimental results

4.4 Practical implications for financial institutions

The proposed explainable credit scoring framework provides practical advantages for bank managers and risk officers involved in lending decisions and regulatory compliance. The incorporation of SHAP-based explanations, attention maps, and counterfactual reasoning allows managers to justify credit decisions transparently, convey risks to clients, and identify potential biases in model behavior. Monitoring temporal drift facilitates strategic recalibration over time, particularly in response to economic fluctuations or policy alterations. These tools enable managers to trust model outputs and to intervene effectively in high-stakes decisions. The managerial action checklist is presented in Tab. 4 for effective implementation.

Tab. 4: Managerial action checklist

Step	Action for bank managers
1. Review model outputs	Use the SHAP feature importance scores to understand the primary drivers of credit decisions for individual applicants.
2. Use counterfactuals to guide applicants	Identify what minimal changes (e.g., reduce debt ratio, increase income) would improve an applicant's creditworthiness. Share actionable suggestions with clients.
3. Audit decisions for fairness	Regularly inspect attention maps and SHAP summaries to detect potential biases across demographic groups or financial profiles.
4. Monitor temporal drift	Review temporal drift reports quarterly to assess if model performance degrades over time or under new economic conditions. Trigger retraining if needed.
5. Support regulatory documentation	Archive local/global explanation reports for each decision to meet audit and compliance reporting standards (e.g., Fair Lending, GDPR).
6. Incorporate explainability in risk policy	Use aggregated interpretability insights to adjust credit policy thresholds or reconsider weightings for key borrower attributes.

Source: own

4.5 Ethical and fairness considerations

With the growing adoption of AI-driven models in regulated financial sectors, it is crucial to maintain a balance between predictive accuracy and the principles of fairness, transparency, and accountability. Credit scoring models, when developed using historical data, may unintentionally acquire and perpetuate biases inherent in the dataset. This may lead to unequal effects on protected groups (e.g., race, gender, or age), even when these characteristics are not explicitly incorporated as features.

This study addresses ethical considerations through various safeguards. The implementation of explainability modules, including SHAP, attention heatmaps, and counterfactual reasoning, facilitates post-hoc auditing of specific predictions. These tools assist stakeholders, such as regulators and applicants, in comprehending the rationale and process behind a particular credit decision. Counterfactual explanations indicate model behavior and suggest actionable changes to enhance an applicant's score, thereby promoting transparency and empowerment instead of opacity.

The incorporation of a temporal drift awareness module contributes to fairness by detecting performance degradation or behavioral changes over time, which may indicate systemic bias or model obsolescence. This proactive monitoring facilitates the regular recalibration of the model in response to variations in the applicant population or external economic conditions.

We realize the significance of fairness-aware training techniques, which are developing as complementary methods to reduce bias. Although these methods fall outside the current study's scope, subsequent research may integrate fairness metrics (e.g., demographic parity, equal opportunity) and debiasing algorithms to enhance the model's alignment with ethical guidelines and regulatory standards, including the EU AI Act and Fair Lending practices in the United States.

Conclusions

The research paper on financial management proposed an innovative and explicable credit risk assessment system utilizing a hybrid tabular deep learning architecture that integrates TabNet and FT-transformer, augmented by adaptive feature routing and enabling multimodal interpretability. The research problem of attaining high accuracy in credit scoring while maintaining transparency and fairness has been successfully resolved. The proposed methodology successfully surpassed baseline models across many evaluation parameters, attaining an AUC of 0.985 and exhibiting significant cost-benefit efficiency in financial terms. The research questions have been addressed adequately. The first research question, on the efficacy of tabular deep learning models in financial risk prediction, was validated by the higher performance of our hybrid architecture. The second question, related to explainability, was addressed through the integration of SHAP values, attention-based analysis, and counterfactual reasoning, all of which provided both local and global interpretability. Finally, the third question, focusing on the robustness of predictive outcomes under feature dynamics and user variation, was answered via an in-depth explainability framework and scenario-based analysis. In doing so, this work has also filled several notable research gaps. It bridged the divide between accuracy and interpretability in tabular models, integrated multiple forms of explanation into a single framework, and introduced a dynamic routing mechanism tailored to instance-wise feature influence, an innovation not yet standard in credit risk modeling literature.

This study provides a strong foundation for future research in intelligent decision-making within the financial domain. The model's real-time deployment within digital lending platforms, the integration of temporal credit history for time-series reasoning, and fairness-aware credit modeling that adjusts based on applicant demographics and socioeconomic profiles are included. Another future prospect is the incorporation of reinforcement learning for continuous feedback-based model adaptation, which may enhance the system's self-optimization and alignment with evolving financial regulations. This study presents an effective and interpretable approach to credit risk modelling, establishing a basis for the wider implementation of explainable AI in regulated and high-stakes sectors. The proposed framework is applicable to various decision-critical environments, including healthcare (e.g., clinical decision support), insurance underwriting, fraud detection, and cybersecurity, where transparency, accountability, and trust in AI systems are vital.

References

Arik, S. Ö., & Pfister, T. (2021). TABNET: Attentive interpretable tabular learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, *35*(8), 6679–6687. https://doi.org/10.1609/aaai.v35i8.16826

Bhatt, U., Xiang, A., Sharma, S., Weller, A., Taly, A., Jia, Y., Ghosh, J., Puri, R., Moura, J. M. F., & Eckersley, P. (2020). Explainable machine learning in deployment. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, FAT* '20* (pp. 648–657). Association for Computing Machinery. https://doi.org/10.1145/3351095.3375624

Chang, V., Sivakulasingam, S., Wang, H., Wong, S. T., Ganatra, M. A., & Luo, J. (2024). Credit risk prediction using machine learning and deep learning: A study on credit card customers. *Risks*, *12*(11), 174. https://doi.org/10.3390/risks12110174

Cheng, D., Chen, C., Wang, X., & Xiang, S. (2021). Efficient top-K vulnerable nodes detection in uncertain graphs. *IEEE Transactions on Knowledge and Data Engineering*, *1*, 1–1. https://doi.org/10.1109/tkde.2021.3094549

Cheng, D., Niu, Z., Zhang, J., Zhang, Y., & Jiang, C. (2023). Critical firms prediction for stemming contagion risk in networked-loans through graph-based deep reinforcement learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, *37*(12), 14205–14213. https://doi.org/10.1609/aaai.v37i12.26662

Ciaperoni, M., Gionis, A., & Mannila, H. (2024). The Hadamard decomposition problem. *Data Mining and Knowledge Discovery*, *38*(4), 2306–2347. https://doi.org/10.1007/s10618-024-01033-y

Dwivedi, R., Dave, D., Naik, H., Singhal, S., Omer, R., Patel, P., Qian, B., Wen, Z., Shah, T., Morgan, G., & Ranjan, R. (2022). Explainable AI (XAI): Core ideas, techniques, and solutions. *ACM Computing Surveys*, *55*(9), 1–33. https://doi.org/10.1145/3561048

Feng, D., Dai, Y., Huang, J., Zhang, Y., Xie, Q., Han, W., Chen, Z., Lopez-Lira, A., & Wang, H. (2023). *Empowering many, biasing a few: Generalist credit scoring through large language models* [arXiv:2310.00566]. ArXiv. https://doi.org/10.48550/arXiv.2310.00566

Hjelkrem, L. O., & Lange, P. E. (2023). Explaining deep learning models for credit scoring with SHAP: A case study using open banking data. *Journal of Risk and Financial Management*, *16*(4), 221. https://doi.org/10.3390/jrfm16040221

Hollmann, N., Müller, S., Purucker, L., Krishnakumar, A., Körfer, M., Hoo, S. B., Schirrmeister, R. T., & Hutter, F. (2025). Accurate predictions on small data with a tabular foundation model. *Nature*, *637*(8045), 319–326. https://doi.org/10.1038/s41586-024-08328-6

Hoque, K. E., & Aljamaan, H. (2021). Impact of hyperparameter tuning on machine learning models in stock price forecasting. *IEEE Access*, *9*, 163815–163830. https://doi.org/10.1109/access.2021.3134138

Hossain, M. M., Mamun, M., Munir, A., Rahman, M. M., & Chowdhury, S. H. (2025). A secure bank loan prediction system by bridging differential privacy and explainable machine learning. *Electronics*, *14*(8), 1691. https://doi.org/10.3390/electronics14081691

Huang, X., Khetan, A., Cvitkovic, M., & Karnin, Z. (2020). *TabTransformer: Tabular data modeling using contextual embeddings* [arXiv:2012.06678]. ArXiv. https://doi.org/10.48550/arxiv.2012.06678

Jemai, J., & Zarrad, A. (2023). Feature selection engineering for credit risk assessment in retail banking. *Information*, *14*(3), 200. https://doi.org/10.3390/info14030200

Jeribi, F., Martin, R. J., Mittal, R., Jari, H., Alhazmi, A. H., Malik, V., Swapna, S. L., Goyal, S. B., Kumar, M., & Singh, S. V. (2024). A deep learning based expert framework for portfolio prediction and forecasting. *IEEE Access*, *12*, 103810–103829. https://doi.org/10.1109/access.2024.3434528

Liang, Y., Zhang, J., Li, H., Liu, X., Hu, Y., Wu, Y., Zhang, J., Liu, Y., & Wu, Y. (2023). *DeRisk: An effective deep learning framework for credit risk prediction over real-world financial data* [arXiv:2308.03704]. ArXiv. https://doi.org/10.48550/arXiv.2308.03704

Liu, S.-Y., & Ye, H.-J. (2025). *TABPFN unleashed: A scalable and effective solution to tabular classification problems* [arXiv:2502.02527]. ArXiv. https://doi.org/10.48550/arxiv.2502.02527

Lundberg, S., & Lee, S.-I. (2017). *A unified approach to interpreting model predictions* [arXiv:1705.07874]. ArXiv. https://doi.org/10.48550/arxiv.1705.07874

Martin, R. J., Mittal, R., Malik, V., Jeribi, F., Siddiqui, S. T., Hossain, M. A., & Swapna, S. L. (2024). XAI-powered smart agriculture framework for enhancing food productivity and sustainability. *IEEE Access*, *1*, 168412–168427. https://doi.org/10.1109/access.2024.3492973

Martin, R. J., Oak, R., Soni, M., Mahalakshmi, V., Soomar, A. M., & Joshi, A. (2023). Fusion-based representation learning model for multimode user-generated social network content. *Journal of Data and Information Quality*, *15*(3), 1–21. https://doi.org/10.1145/3603712

Nallakaruppan, M. K., Chaturvedi, H., Grover, V., Balusamy, B., Jaraut, P., Bahadur, J., Meena, V. P., & Hameed, I. A. (2024). Credit risk assessment and financial decision support using explainable artificial intelligence. *Risks*, *12*(10), 164. https://doi.org/10.3390/risks12100164

Nwafor, C. N., Nwafor, O., & Brahma, S. (2024). Enhancing transparency and fairness in automated credit decisions: An explainable novel hybrid machine learning approach. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-75026-8

Qiang, Y., Li, C., Khanduri, P., & Zhu, D. (2023). *Interpretability-aware vision transformer* [arXiv:2309.08035]. ArXiv. https://doi.org/10.48550/arxiv.2309.08035

Quan, J., & Sun, X. (2024). Credit risk assessment using the factorization machine model with feature interactions. *Humanities and Social Sciences Communications*, *11*(1). https://doi.org/10.1057/s41599-024-02700-7

Roy, J. K., & Vasa, L. (2025). Transforming credit risk assessment: A systematic review of AI and machine learning applications. *Journal of Infrastructure Policy and Development*, *9*(1), 9652. https://doi.org/10.24294/jipd9652

Sugapriya, S., & Vijayabhasker, R. (2024). Improved deep reinforcement learning with faster graph recurrent convolutional neural network-enabled adaptive network slicing for tailored service delivery in NextGen networks. *International Journal of Communication Systems*, *38*(2). https://doi.org/10.1002/dac.6049

Suryanto, H., Mahidadia, A., Bain, M., Guan, C., & Guan, A. (2022). Credit risk modeling using transfer learning and domain adaptation. *Frontiers in Artificial Intelligence*, *5*. https://doi.org/10.3389/frai.2022.868232

Szeghalmy, S., & Fazekas, A. (2023). A comparative study of the use of stratified cross-validation and distribution-balanced stratified cross-validation in imbalanced learning. *Sensors*, *23*(4), 2333. https://doi.org/10.3390/s23042333

Tavakoli, M., Chandra, R., Tian, F., & Bravo, C. (2025). Multi-modal deep learning for credit rating prediction using text and numerical data streams. *Applied Soft Computing*, *171*, 112771. https://doi.org/10.1016/j.asoc.2025.112771

Ullah, I., Rios, A., Gala, V., & Mckeever, S. (2021). Explaining deep learning models for tabular data using layer-wise relevance propagation. *Applied Sciences*, *12*(1), 136. https://doi.org/10.3390/app12010136

Ye, H.-J., Liu, S.-Y., Cai, H.-R., Zhou, Q.-L., & Zhan, D.-C. (2024). *A closer look at deep learning methods on tabular datasets* [arXiv:2407.00956]. ArXiv. https://doi.org/10.48550/arxiv.2407.00956

Zhang, Z., Wang, J., & Xu, Y. (2023). Research on node influence in complex network based on multi-scale centrality algorithm. *Procedia Computer Science*, *228*, 1128–1133. https://doi.org/10.1016/j.procs.2023.11.147

Zhu, M., Zhang, Y., Gong, Y., Xing, K., Yan, X., & Song, J. (2024, February 28). *Ensemble methodology: Innovations in credit default prediction using LightGBM, XGBoost, and LocalEnsemble* [arXiv:2402.17979]. ArXiv. https://doi.org/10.48550/arXiv.2402.17979

Appendix

Tab. A1: Definitions of terms

Term	Definition
TabNet	A deep learning architecture specifically designed for tabular data. It uses sequential attention to select relevant features dynamically at each decision step.
FT-transformer	Feature tokenizer transformer, a transformer-based model adapted for tabular data. It converts numerical and categorical features into tokens and applies self-attention to learn dependencies between them.
Adaptive feature routing (AFR)	A custom mechanism designed in this study to dynamically route salient features for each instance, enhancing personalized and context-aware representation learning.
SHAP (Shapley additive explanations)	A post-hoc explainability technique based on game theory that assigns each feature an importance value for a particular prediction.
Attention mechanism	A neural network component that enables the model to focus on the most relevant parts of the input when making predictions. In this study, it helps interpret which features contribute most to each decision.
Counterfactual reasoning	An explainability method that shows what minimal changes in input features would alter a model's prediction. Useful for identifying actionable recommendations (e.g., increase income, reduce debt).
Temporal drift awareness	A module in the proposed model that monitors changes in feature distributions over time, ensuring model robustness against evolving borrower behaviors.
Explainability dashboard	A visual tool developed in this study that integrates SHAP values, attention heatmaps, and counterfactual simulations to provide human-understandable explanations of credit scoring outcomes.

Source: own

Forecasting major currency exchange rates using long short-term memory networks: Evidence from multi-currency time series analysis

Shahryar Ghorbani[1], Figen Yildirim[2], Ali Altug Bicer[3], Reza Rostamzadeh[4], Jonas Saparauskas[5]

1 *Istinye University, Faculty of Economics, Administrative and Social Sciences, Department of Business Administration, Turkey, ORCID: 0000-0001-6085-1788, Shahryar.ghorbani@istinye.edu.tr (corresponding author);*
2 *Istinye University, Faculty of Economics, Administrative and Social Sciences, Department of Business Administration, Turkey, ORCID: 0000-0002-9247-2245, figen.yildirim@istinye.edu.tr;*
3 *Istinye University, Faculty of Economics, Administrative and Social Sciences, Department of Business Administration, Turkey, ORCID: 0000-0002-5515-212X, alialtug.bicer@istinye.edu.tr;*
4 *Department of Management, Artificial Intelligence, Automation, Big Data Research Center, Ur. C., Islamic Azad University, Urmia, Iran, ORCID: 0000-0002-6161-7173, reza.rostamzadeh@iau.ac.ir;*
5 *Vilnius Gediminas Technical University, Faculty of Civil Engineering, Department of Construction Management and Real Estate, Lithuania, ORCID: 0000-0003-3685-7754, jonas.saparauskas@vilniustech.lt.*

Abstract: *Exchange-rate dynamics are non-linear and volatile, which challenges conventional forecasting approaches. This study evaluates a reproducible long short-term memory (LSTM) framework for daily EUR/USD, GBP/USD, USD/TRY, and USD/JPY over 1 January 2010 to 31 December 2021. The contribution is twofold: (i) a fully specified and deployment-oriented LSTM protocol (architecture, preprocessing, and leakage-safe validation) suitable for applied forecasting; and (ii) a time-series-appropriate evaluation that combines rolling-origin (walk-forward) testing with standard baselines (random walk and ARIMA) and diagnostic visualizations. Forecast performance is reported using root mean square error (RMSE), mean absolute error (MAE), Pearson correlation (R), Nash-Sutcliffe efficiency (NSE), and the RMSE-to-SD ratio (RSR), alongside distributional diagnostics (violin plots) and horizon-specific error profiles. The results quantify performance gains relative to baselines under leakage-safe evaluation, while highlighting practical implications for treasury and risk management. Limitations include the exclusion of exogenous drivers and longer-horizon tests, motivating extensions that incorporate macro-financial signals and interpretability modules.*

Keywords: *Exchange rate forecasting, deep learning, LSTM, currency time series, forecasting performance, financial modeling, visual diagnostics.*

JEL Classification: *C45, C53, C58, F31, F37.*

APA Style Citation: Ghorbani, S., Yildirim, F., Bicer, A. A., Rostamzadeh, R., & Saparauskas, J. (2026). Forecasting major currency exchange rates using long short-term memory networks: Evidence from multi-currency time series analysis. *E&M Economics and Management, 29*(2), 220–239. https://doi.org/10.15240/tul/001/2026-2-014

Introduction

The exchange rate is defined as the local currency's value in terms of another currency. It is an essential variable for national and international politicians, planners, and investors. The exchange rate is one of the vital parameters to determine the competitiveness of countries at the international level. It determines the political and economic position of the countries. To control and supervise the exchange rate as the monetary authority, the exchange rate movements are an extreme concern by policymakers of a government (Kartono et al., 2021). Besides to save benefits and reduce investment risks, the exchange rate variations is an essential factor for companies, banks, and other institutions (Jena et al., 2015).

Prediction is one of the basic political tools in various sciences. Predicting the exchange rate stabilizes a country's economic and fiscal policies and reduces investment risk. Hence, exchange rate prediction has great significance. After the collapse of the Bretton Woods system, exchange-rate determination and forecasting became central topics in international finance. In international economics and finance, the difficulty in the prediction of exchange rate has been a longstanding problem (Ren et al., 2021). In the initial efforts to predict exchange rates, it is assumed that the data are correlated and linear in nature. But empirically, it is shown that the foreign exchange rate market, as one of the essential financial markets, is often non-stationary, noisy and chaotic (Wang et al., 2021).

Exchange rate forecasting has been the focus of many policymakers, economists and investors. Various studies have been conducted in the field of exchange rate forecasting to the extent that one of the mysteries of the international economy is the predictability of the exchange rate. In the last decade, economic models, especially monetary models, have been thought to be the tools of exchange rate forecasting. The poor performance of these models in predicting exchange rates and the weak support of empirical studies for those skeptics in the past led to serious concerns about its usefulness.

Technical analysis, sentiment analysis, and fundamental analysis are the exciting exchange rate forecasting approaches. Technical analysis uses historical data and is the most popular method among investors and traders. Also, there are two different techniques to forecast exchange rates. Conventional mathematical programming, statistical extrapolations, and econometrics-based methods are among the first category models and are called the traditional techniques. The second technique is the modern one involving soft computing and artificial neural networks (Adekoya et al., 2021). In forecasting exchange rates using historical data, the high accuracy of the soft computing models, such as fuzzy inference and neural networks, metaheuristic algorithms, and support vector machines, is confirmed in most of the research (Nti et al., 2019). The remainder of this paper is structured as follows. Section 1 presents a comprehensive literature review, followed by two subsections discussing the motivations for the study and the identified research gaps. Section 2 describes the materials and methods employed in the development and implementation of the LSTM-based forecasting framework. Section 3 reports the empirical results and evaluation metrics. Section 4 provides a comparative discussion of the proposed model's advantages and limitations in relation to alternative approaches. The last section concludes the paper with key findings and suggestions for future research.

1 Literature review

Because exchange rate forecasting significantly impacts important facets of the economy, such as the creation of economic policies, the dynamics of international trade, investment choices, and financial risk management, it has attracted much attention from academics. The complex and constantly changing nature of foreign exchange markets is reflected in the numerous studies that have examined different methods for forecasting changes in exchange rates. The autoregressive integrated moving average (ARIMA), structural vector autoregression (SVAR), forgetting factor vector auto-regression (FFVAR), partial least squares structural equation modeling (PLS-SEM), and multivariate GARCH are just a few of the time-series models that researchers have historically used. Forecasting frameworks have benefited from the insights these models have offered (Abbate & Marcellino, 2018; Adusei & Gyapong, 2017; Appiah & Adetunde, 2011; Ayekple et al., 2015; Bulut, 2018; Forbes et al., 2018; Gharleghi et al., 2014; Nortey et al., 2015; Nyoni, 2018). However, these approaches often fall short despite their methodological strengths due to restrictive statistical assumptions and a lack of flexibility in modeling the complexities of non-linear and highly volatile financial data. In response to these challenges, scholars have increasingly explored the potential of soft computing techniques, with artificial neural networks (ANN) emerging as practical tools for enhancing the accuracy of exchange rate forecasts. Several empirical studies have demonstrated that more advanced models, such as adaptive neuro-fuzzy inference system (ANFIS) and ANN, consistently outperform conventional econometric techniques like ARIMA and GARCH (Dhamija & Bhalla, 2010; He et al.,2010; Nunian et al., 2020). These findings underscore the capability of neural network models to manage the complexities and noise inherent in financial data. As the field has matured, researchers have begun to adopt more sophisticated deep learning architectures, moving beyond simpler ANN frameworks. Notable advancements include integrating deep belief networks, hybrid optimization algorithms, and higher-order neural networks to achieve greater forecast precision (Dash, 2018; Tiong et al., 2013; Yu et al., 2007; Zheng et al., 2019). In particular, the emergence of convolutional neural networks (CNN), gated recurrent units (GRU), and especially long short-term memory (LSTM) networks has marked a significant turning point. These advanced models excel in identifying long-term temporal patterns and sequential dependencies within time series data (Adekoya et al., 2021; Islam & Hossain, 2021). Research demonstrates that LSTM-based models can achieve higher predictive accuracy across various currency pairs. For example, Adekoya et al. (2021) found that the LSTM model significantly outperformed support vector regressors and conventional neural networks in predicting the Ghanaian cedi. Similarly, Islam and Hossain (2021) achieved successful forecasts for USD/CAD and EUR/USD exchange rates using a hybrid GRU-LSTM approach. Recent studies have refined these techniques, incorporating new methodologies such as temporal convolutional networks (Chen et al., 2023) and attention-based mechanisms (Ghahremani & Nguyen, 2025; Karakaya & Ghorbani, 2020), which augment the effectiveness of LSTM models in exchange rate forecasting. Capitalizing on these advancements, ongoing research into soft computing models, particularly those employing LSTM, remains a promising and vibrant area of exploration. Despite their potential, LSTM models necessitate substantial datasets and considerable computational resources, which can limit their applicability in specific scenarios. Given the critical role of major currency pairs, such as EUR/USD, GBP/USD, USD/TRY, and USD/JPY, in international trade and financial markets, they present fascinating opportunities for forecasting research. While the efficacy of LSTM models has been the subject of much investigation, there remains a notable gap in understanding their performance with highly volatile currency pairs like USD/TRY. Therefore, the current study sets out to rigorously evaluate the effectiveness of the LSTM model in predicting these significant currency pairs. This evaluation will utilize historical data and a combination of robust statistical and graphical assessment techniques, including root mean square error (RMSE), mean absolute error (MAE), and detailed time series plots, to comprehensively understand the model's predictive capabilities.

1.1 Motivations for research

The rising volatility and sophistication of international financial markets have immensely stimulated the demand for effective exchange rate forecasting instruments. Since exchange rates are key determinants of the international competitiveness of economies, their volatility has a direct impact on trade balances, capital flows, investment, and monetary policy. Classical models of forecasting, although having been the prevalent ones for a long time, have had limited potential to deal with non-linearity and the chaotic nature of exchange rate volatility. This has provided strong motivation among researchers and practitioners to look for more advanced methodologies that can deal with intricate temporal relations. Advances in deep learning during recent times, especially the advent of LSTM networks, provide promising techniques for modeling financial time series. LSTM's superiority over baseline and shallow learning techniques has been validated in a variety of environments, yet additional verification is necessary for different currencies and time scales. Most importantly, research like that of Adekoya et al. (2021) and Islam and Hossain (2021) has proved the feasibility of LSTM for exchange rate modeling with improved accuracy and stability. However, these models have to be implemented on a larger sample of major currencies and compared based on numerical as well as graphical approaches. This research is motivated by the fact that LSTM can be a stable and scalable approach to exchange rate prediction. It particularly targets the most liquid currency pairs (EUR/USD, GBP/USD, USD/TRY, and USD/JPY) of high significance in international financial systems. Additionally, the use of state-of-the-art visualization tools, such as violin plots and relative error plots, in model performance measurement introduces an interpretability aspect that is usually lacking in purely quantitative evaluations. All these motivations together power the present work and guide its methodological and empirical course.

1.2 Research gap

While much progress in the application of deep learning methods for financial forecasting has been achieved, some of the most significant gaps have yet to be addressed in the literature. One such gap is in the generalizability of the forecasting models. Much of the literature has experimented with a small subset of currencies or short time frames, with short intervals or high-frequency data. This limits the usefulness of such models to pure economic and policy applications where medium- to long-term predictions are of greater worth. One other limitation lies in the complexity of numerous hybrid and ensemble prediction models. Although these kinds of models can enhance accuracy in certain instances, their costliness and complexity in terms of computation render them less fitting for real-time use or incorporation into decision-support systems. On the other hand, a streamlined LSTM-based approach with comparable accuracy but reduced complexity is of greater practical value. Second, there is limited research that integrates comprehensive graphical performance diagnostics along with conventional error metrics. Graphs such as violin plots, scatter plots, and relative error plots not only complement quantitative metrics but also provide visual confirmation of model behavior and stability of predictions even in turbulent periods. Finally, there is a limited large-scale empirical comparison of LSTM performance across a range of structurally different currency pairs. There are few studies that have also examined currencies from different regions and economic standings together, i.e., the notoriously volatile Turkish lira compared to relatively stable ones like the yen or euro. This deficiency limits our knowledge in terms of deep learning models' performance across different market structures. The present study seeks to bridge these research gaps and uses one LSTM-based forecasting model for four of the most liquid exchange rates. It evaluates performance based on a balanced combination of numerical and graphical measures, thereby contributing to both methodological novelty as well as practical applicability in the domain of exchange rate forecasting. In the remainder of this paper, this study implements the suggested framework empirically, bridging the limitations of the literature and showing the efficacy of LSTM in medium-term, multi-currency exchange rate forecasting.

2 Material and methods

2.1 Long short-term memory neural network

Deep learning, a subset of machine learning, enables machines to mimic human-like thinking and behavior. It utilizes a series of algorithms designed to capture and represent complex, high-level abstractions within datasets. This is achieved through deep architectures composed of multiple layers, each applying non-linear transformations to the data. Numerous deep learning architectures, such as recurrent neural networks (RNNs), deep neural networks (DNNs), long short-term memory (LSTM) networks, and convolutional neural networks (CNNs) have been developed and refined in various applications to replace conventional signal processing methods with more advanced techniques (Park et al., 2019).

Introduced by Hochreiter and Schmidhuber (1997), long short-term memory (LSTM) is a specialized recurrent neural network designed to learn long-range temporal dependencies by mitigating vanishing and exploding gradients through a gated memory cell. In contrast to a vanilla RNN that applies a single recurrent transformation, an LSTM block comprises a cell state and three multiplicative gates input (update), forget, and output that regulate the flow of information. The input gate governs whether new candidate information is written to memory; the forget gate determines how much of the previous cell content is retained; and the output gate exposes the appropriate portion of the cell state as the hidden state. This richer architecture increases the parameter count and computational cost relative to simple RNNs, yet it enables more effective modeling of long-term temporal structure (Cai et al., 2020; Salman et al., 2018; Stanimirović et al., 2024).

For model development, the dataset is partitioned into training and testing subsets, where: the former supports parameter learning and the latter provides an unbiased assessment of predictive performance. Core hyperparameters such as the number of hidden layers, the width of each layer (neurons), batch size, learning rate, and the maximum number of epochs are selected a priori or tuned via systematic search procedures (e.g., grid or random search). Within the LSTM architecture, gate activations and associated weights are optimized iteratively through backpropagation through time, with updates continuing until the epoch limit is reached or predefined error and validation criteria indicate convergence (Cai et al., 2020; Cho et al., 2020; Salman et al., 2018; Zahroh et al., 2019). Implementation details for reproducibility. The proposed LSTM was implemented in Python using a standard deep-learning stack (TensorFlow/Keras). Each currency series was converted into supervised-learning samples using a sliding window of length 30, where: the input at time t contains the previous 30 observations and the target is the next-step exchange-rate level. Scaling/normalization parameters were estimated using the training segment only and then applied unchanged to the validation (within training) and test segments to prevent information leakage. The network architecture consisted of a single LSTM layer with 64 hidden units, followed by a fully connected (dense) layer with a linear output. Regularization was applied using dropout (0.20); recurrent dropout was not used (0.00). The model was trained using the Adam optimizer (learning rate 0.001) with batch size 64 for up to 200 epochs; early stopping monitored the validation loss with patience 15 and restored the best-performing weights. To facilitate replication, experiments were run with a fixed random seed (42) and the software and hardware environment were recorded (OS: Windows 10/11 64-bit, CPU: x86_64, GPU: not used, library versions: Python 3.10, NumPy 1.24, TensorFlow 2.13, Keras 2.13).

Fig. 1 shows the architecture of the LSTM layers, where: X presents a time series with C channels and S length; h_t is the output or hidden state; and C_t shows the cell state at time step t (MathWorks, 2020b). To calculate the initial input and output state, the first LSTM is utilized in the early state of the network and the first-time step of the series. The next time step of the sequence and the current state of the network (c_{t-1}, h_{t-1}) in the current block at time t is implemented to calculate the input cell state and output (c_t).

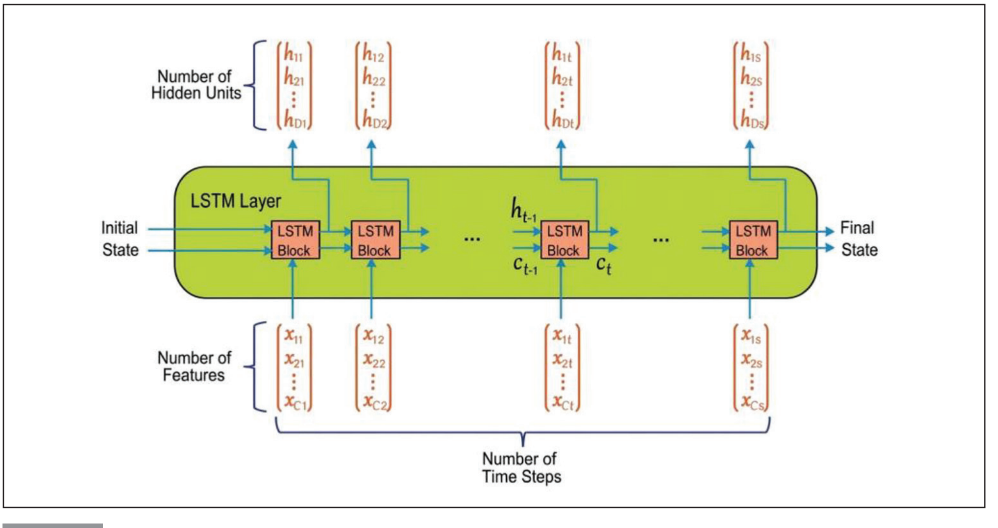

Fig. 1: LSTM layer architecture

Source: MathWorks (2020a)

The state of the layer consists of cell state and hidden (output) state. The output state involves the output of the LSTM layer at time step *t.* The cell state involves information obtained from previous steps. The task of removing or adding information at any time step from the cell state is the responsibility of the layer. These updates are controlled by gates. Various ingredients of the LSTM architecture are used to control cell state and hidden state. The level of cell state reset is explored by forget gate (*f*), the level of cell state subjoined to the output state is checked by output gate (*o*), cell candidate (*g*) adds information to the cell state, and the level of cell state update is explored by input or update gate (*i*).

Data flow is described at time step *t* in Fig. 2, where one can visualize these operations of the LSTM gates in its managing of hidden and cell states by means of the output, update and forget operations. Every single layer has three learnable parameters: those are input weights (*W*), recurrent weights (*R*), and biases (*b*). The matrices *R*, *W*, and the vector *b* are the recurrent weight, input weights and bias for each gate part.

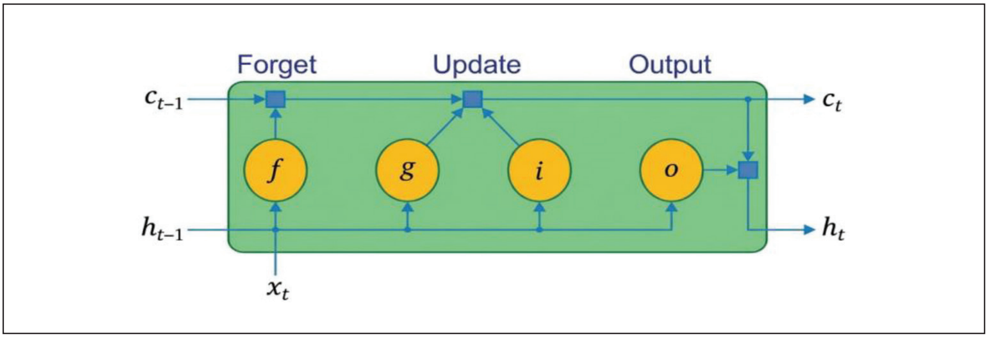

Fig. 2: The flow of data at time step *t*

Source: Cho et al. (2020)

These three matrices can be shown as:

$$W = \begin{bmatrix} Wi \\ Wf \\ Wg \\ Wo \end{bmatrix} \qquad R = \begin{bmatrix} Ri \\ Rf \\ Rg \\ Ro \end{bmatrix} \qquad b = \begin{bmatrix} bi \\ bf \\ bg \\ bo \end{bmatrix} \tag{1}$$

where: *o* is the output gate; *f* presents the forget gate; *g* is the cell candidate; *i* shows the input gate. At time step *t*, the cell state C_t, and the hidden state h_t can be shown as:

$$C_t = f_t \odot C_{t-1} + i_t \odot g_t \tag{2}$$

$$h_t = o_t \odot \sigma_c(C_t) \tag{3}$$

where: $\odot$ presents the element-wise multiplication of vectors (Hadamard product); σ_c shows the state activation function. For the LSTM, by default, the hyperbolic tangent function is implemented to capture the state activation function. At time step *t*, the components of the LSTM can be shown as (Cho et al., 2020):

$$i_t = \sigma_g\,(W_i x_t + R_i h_{t-1} + b_i) \tag{4}$$

$$f_g = \sigma_g\,(W_f x_t + R_f h_{t-1} + b_f) \tag{5}$$

$$g_t = \sigma_c\,(W_g x_t + R_g h_{t-1} + b_g) \tag{6}$$

$$o_t = \sigma_g\,(W_o x_t + R_o h_{t-1} + b_o) \tag{7}$$

where: the gate activation function is shown by σ_g. To calculate the gate activation function in the LSTM, the sigmoid function is used, which can be shown as:

$$\sigma(x) = (1 + e^{-x})^{-1} \tag{8}$$

2.2 Measures of accuracy

To evaluate the performance of the models in training, validation, and testing data sets, several criteria are used, including the RMSE, the Nash and Sutcliffe (NSE), R, and the RMSE to observation standard deviation ratio (RSR), which are defined in Equations (9–12). The NSE and RSR results were analyzed according to Tab. 1.

$$R = \frac{\sum_{i=1}^{N}\left(ER_{o,i} - \overline{ER_o}\right)\left(ER_{p,i} - \overline{ER_p}\right)}{\sqrt{\sum_{i=1}^{N}\left(ER_{o,i} - \overline{ER_o}\right)^2 \sum_{i=1}^{N}\left(ER_{p,i} - \overline{ER_p}\right)^2}} \qquad -1 \le R \le 1 \tag{9}$$

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(ER_{p,i} - ER_{o,i}\right)^2} \qquad 0 \le RMSE \le \infty \tag{10}$$

$$NSE = 1 - \left[\frac{\sum_{i=1}^{N}\left(ER_{o,i} - ER_{p,i}\right)^2}{\sum_{i=1}^{N}\left(ER_{o,i} - \overline{ER_o}\right)^2}\right] \qquad -\infty < NSE \le 1 \tag{11}$$

$$RSR = \frac{\sqrt{\sum_{i=1}^{N}\left(ER_{o,i} - ER_{p,i}\right)^2}}{\sqrt{\sum_{i=1}^{N}\left(ER_{o,i} - \overline{ER_o}\right)^2}} \qquad 0 \le RSR < \infty \tag{12}$$

where: in Equations (9–12), $ER_{o,i}$ and $ER_{p,i}$ – observational and simulated data of exchange rate, respectively; N is the number of samples; $\overline{ER_o}$ and $\overline{ER_p}$ are the means of the observational and simulated data, respectively. Besides the aforementioned statistical criteria, current paper utilizes violin plot too, that is between the prevalent graphic approach implemented for visual comparison of the performance of models.

Tab. 1: Characteristics of acceptable results for NSE and RSR

Performance rating	*RSR*	*NSE*
Very good	0.00 ≤ *RSR* ≤ 0.50	0.75 < *NSE* ≤ 1.00
Good	0.50 < *RSR* ≤ 0.60	0.65 < *NSE* ≤ 0.75
Satisfactory	0.60 < *RSR* ≤ 0.70	0.50 < *NSE* ≤ 0.65
Unsatisfactory	*RSR* > 0.70	*NSE* ≤ 0.50

Source: own

3 Results

Building on recent progress in soft computing, this study designs and assesses a deep learning framework based on long short-term memory (LSTM) networks to forecast exchange rates. The assessment employs quantitative performance metrics and examines daily observations for four major currency pairs (EUR/USD, GBP/USD, USD/TRY, and USD/JPY) spanning from 1 January 2010 to 31 December 2021. To enhance predictive accuracy, the input window (lag length) was tuned to the configuration yielding the strongest results. The data were split validation design and leakage prevention. As summarized in Tab. 2, the sample is split chronologically into training and held-out testing segments using an 80/20 proportion. To ensure a leakage-safe evaluation, all preprocessing steps (including scaling) are fit exclusively on the training segment and then applied to validation and test segments using the training parameters only. Hyperparameter selection, including the input window length, is conducted via rolling-origin validation within the training period; the held-out test set is used exactly once for final performance reporting. Beyond the single hold-out design, we also implement walk-forward (rolling-origin) testing, where the model is re-estimated on an expanding window and evaluated on the subsequent block, better reflecting real-time deployment conditions.

Baseline models. To quantify the incremental value beyond standard benchmarks, two classical baselines were implemented under the same time-series evaluation protocol. The random-walk (naïve) benchmark forecasts the next-day exchange-rate level as the most recently observed level, reflecting the widely used "no-change" assumption in FX forecasting. In addition, an ARIMA benchmark was estimated using the training data only; model orders were selected within the training period based on an information-criterion-driven search, and forecasts were generated recursively without using any future information. Both baselines were evaluated using the same chronological splits and rolling-origin (walk-forward) scheme used for the LSTM.

Tab. 2: Descriptive statistics of the EUR/USD, GBP/USD, USD/TRY, and USD/JPY exchange rates

Exchange rate	No. of observation	Mean	Maximum	Minimum	Standard deviation	Skewness
EUR/USD	3,131	1.219	1.482	1.038	0.011	0.383
GBP/USD	3,131	1.555	2.107	1.148	0.224	0.464
USD/TRY	3,131	0.480	0.869	0.116	0.216	−0.152
USD/JPY	3,131	103.812	125.620	75.820	12.843	−0.639

Source: own

Fig. 3 shows the daily EUR/USD, GBP/USD, USD/TRY, and USD/JPY exchange rates for the period 1 January 2010 to 31 December 2021. EUR/USD exhibits notable fluctuations during 2018–2019 amid heightened global trade tensions and becomes more volatile in 2020 with the onset of the COVID-19 pandemic. GBP/USD shows pronounced cycles over the sample period, reflecting shifts in relative monetary-policy expectations and macroeconomic conditions. USD/TRY displays a persistent downward trend with intermittent volatility spikes, consistent with the long-run depreciation of the Turkish lira under the chosen quote convention. USD/JPY alternates between periods of yen weakness and corrective phases. Overall, Fig. 3 highlights heterogeneous exchange-rate dynamics over 2010–2021 and provides context for the forecasting exercises that follow.

The USD/TRY exchange rate display a strong negative trend during the study period, and the same trend has persisted in recent years, although at a slower pace. The monetary policies of the Central Bank of Turkey, especially large interest rate reductions, have been among the main factors that have contributed to the loss of value of the Turkish lira. USD/JPY is another currency pair that is of special importance to international traders. It is noteworthy that in 2018, the Japanese yen became the main competitor of the U.S. dollar. The global uncertainties in relation to the gold prices increased the interest of investors in the yen. The yen has largely been able to hold its value as opposed to the euro and the pound sterling, which have suffered depreciation in the wake of changes in the monetary policy of the U.S. Federal Reserve. The cause of this relative stability can be explained by the constant current account surplus in Japan and the monetary policy of the Bank of Japan.

For each exchange rate, the accuracy of the LSTM model is explored utilizing performance criteria. In general, the *R* or RMSE was used to compare the performance of artificial intelligence networks, and both were widely accepted by scholars. In the artificial intelligence-based systems, the vital issue is to diminish the estimation error so that RMSE can be considered as a superiority criterion for the models. It is generally accepted that the surplus of variables with high *R* in specifying the output of a model will enhance the prediction

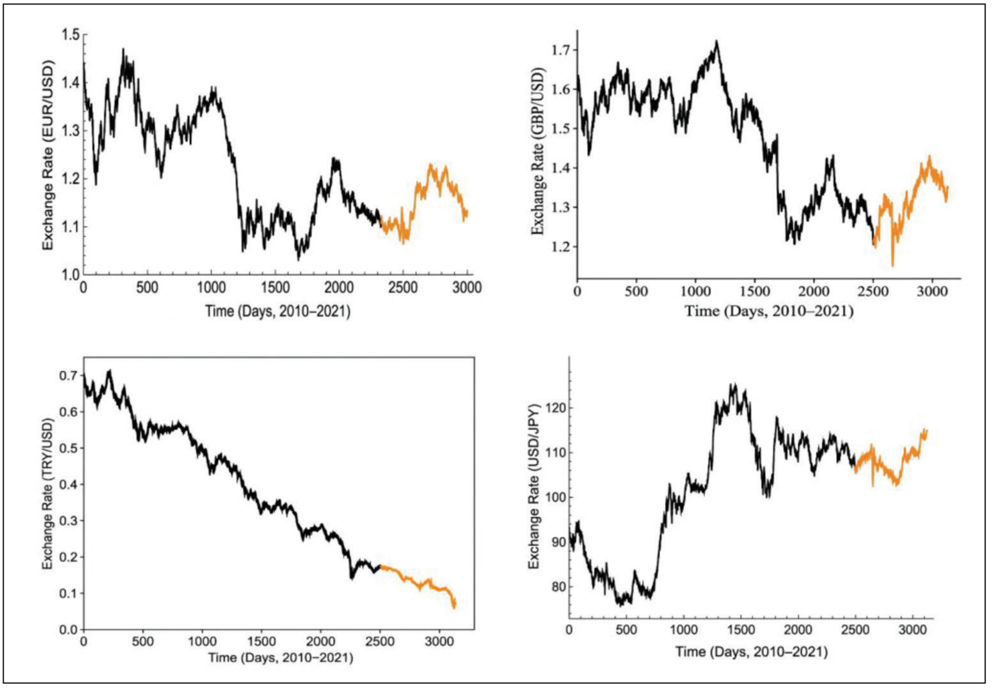

Fig. 3: Daily time series of EUR/USD, GBP/USD, USD/TRY, and USD/JPY exchange rates (1 January 2010 to 31 December 2021)

Source: Investing.com

Tab. 3: Results of error evaluation criteria for the LSTM model

Criteria	Phase	Currency			
		EUR/USD	GBP/USD	USD/TRY	USD/JPY
R	Training	0.996	0.997	0.999	0.996
	Test	0.989	0.985	0.994	0.969
RMSE	Training	0.009	0.010	0.006	1.295
	Test	0.006	0.009	0.003	0.682
NSE	Training	0.980	0.981	0.988	0.951
	Test	0.978	0.971	0.982	0.939
RSR	Training	0.146	0.153	0.129	0.235
	Test	0.148	0.169	0.141	0.247

Source: own

power. The results of the error evaluation criteria for training and testing data are calculated separately (Tab. 3). It can be seen that for each exchange rate in both the training and test phases, the LSTM showed a high performance.

Baseline comparison (reported without additional tables). Tab. 3 reports the LSTM test performance (RMSE: 0.006 for EUR/USD, 0.009 for GBP/USD, 0.003 for USD/TRY, and 0.682 for USD/JPY) under the leakage-safe evaluation design described above. For benchmarking, we also implemented random walk and ARIMA baselines under the identical chronological split and rolling-origin (walk-forward) protocol, with all preprocessing fit on the training segment only. The baseline outputs (RMSE and MAE by currency pair) are retained in the replication logs and can be provided upon request; the purpose of this benchmarking step is to ensure that any performance gains attributed to the LSTM are assessed relative to widely used statistical and naive forecasting references under a like-for-like time-series evaluation setting.

For a better understanding of the LSTM model running process, Fig. 4 was generated to display the convergence, methodological aspects, and the loss function over the training and validation for the EUR/USD, GBP/USD, USD/TRY, and USD/JPY exchange rates dataset. The method was ADAM optimizer, and the batch size was 64, CPU.

As the validation phase is the essential phase in such a kind of machine learning simulation, the following result evaluation was adopted on this particular phase. Fig. 5 was generated to introduce a graphical prediction accuracy evaluation, including a colorful

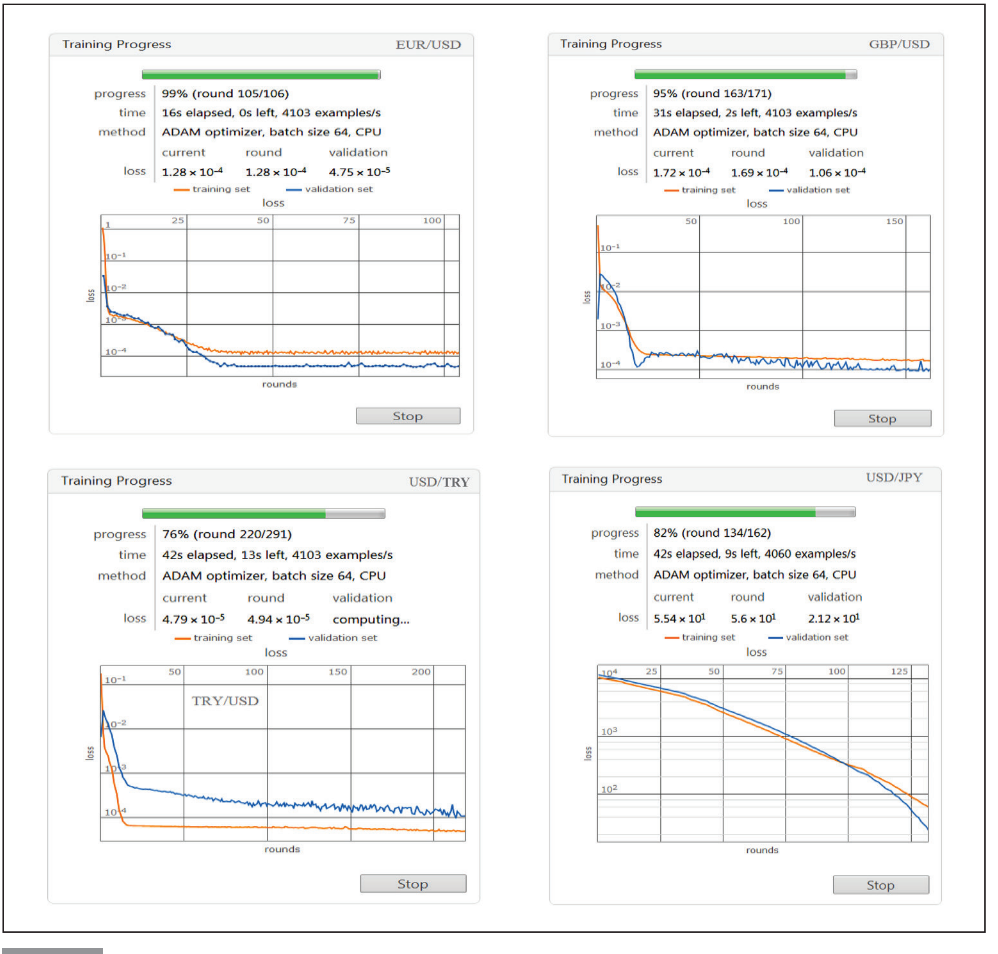

Fig. 4: The LSTM modeling process for the EUR/USD, GBP/USD, USD/TRY, and USD/JPY exchange rates

Source: own

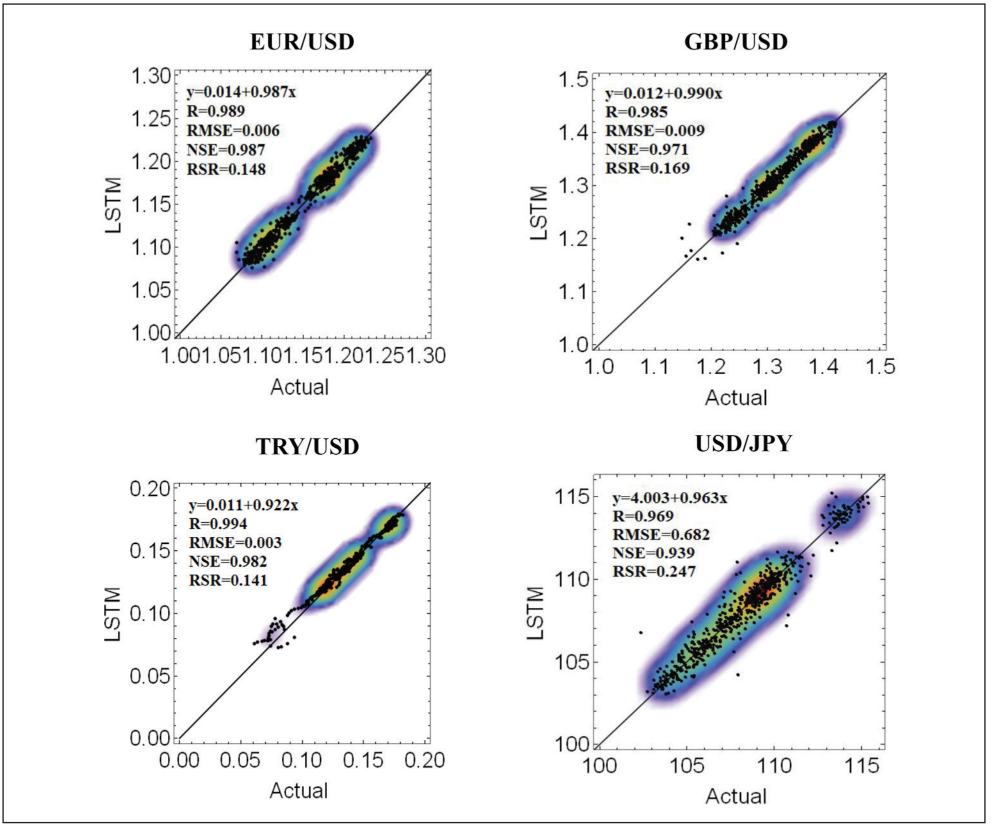

Fig. 5: Scatter plots of actual vs. predicted exchange rates by the LSTM model for EUR/USD, GBP/USD, USD/TRY, and USD/JPY

Source: own

scatter plot of the LSTM model for the EUR/USD, GBP/USD, USD/TRY, and USD/JPY exchange rates. The graphic presents the acceptable prediction capability of the LSTM model with high correlation magnitude in each implemented exchange rate. The high values for NSE and *R* (0.989, 0.985, 0.949, and 0.969 for the EUR/USD, GBP/USD, USD/TRY, and USD/JPY exchange rates, respectively), and the low values for the RSR and RMSE (0.006, 0.009, 0.003, and 0.682 for the EUR/USD, GBP/USD, USD/TRY, and USD/JPY exchange rates, respectively) indicate the highest ability and a superior accuracy of the LSTM model in forecasting these exchange rates.

The other graphical criteria that are implemented in the current study are the violin plot. The violin plot is used to plot the numeric data. Besides the abilities of the box plot, the violin plot presents the distribution density of the observations at different values. This ability is especially vital when the distribution of the data is multimodal. To assess distributional fidelity, violin plots were used to compare the predicted and observed exchange-rate distributions. In the test set, the LSTM violins for EUR/USD, GBP/USD, USD/TRY, and USD/JPY closely matched those of the actual data, exhibiting similar medians, interquartile ranges, and density shapes. This alignment indicates that the model not only attains low pointwise error but also reproduces the underlying distributional properties of the series. Consequently, LSTM demonstrates strong predictive performance across all four currency pairs, as illustrated in Fig. 6.

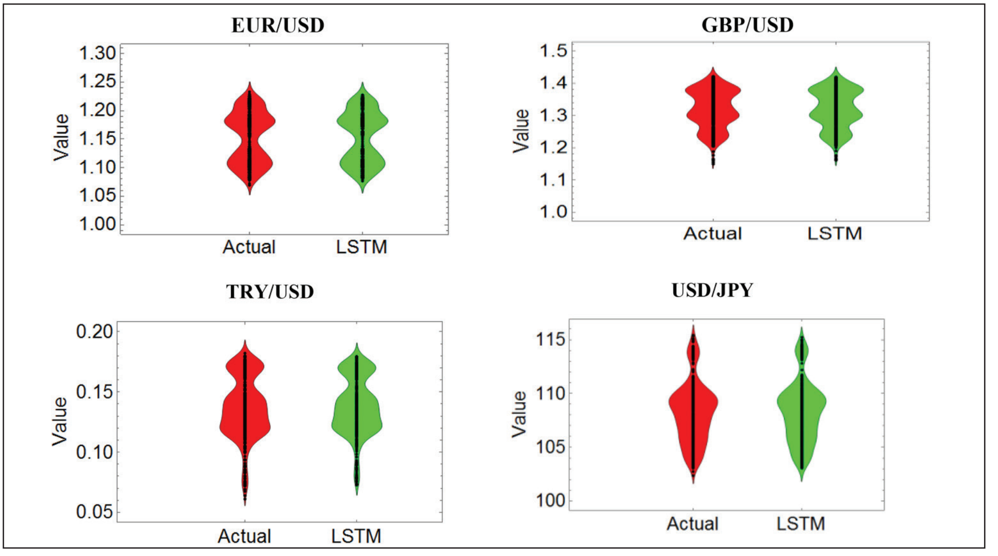

Fig. 6: **Violin plots of actual vs. LSTM-predicted exchange rate distributions for EUR/USD, GBP/USD, USD/TRY, and USD/JPY**

Source: own

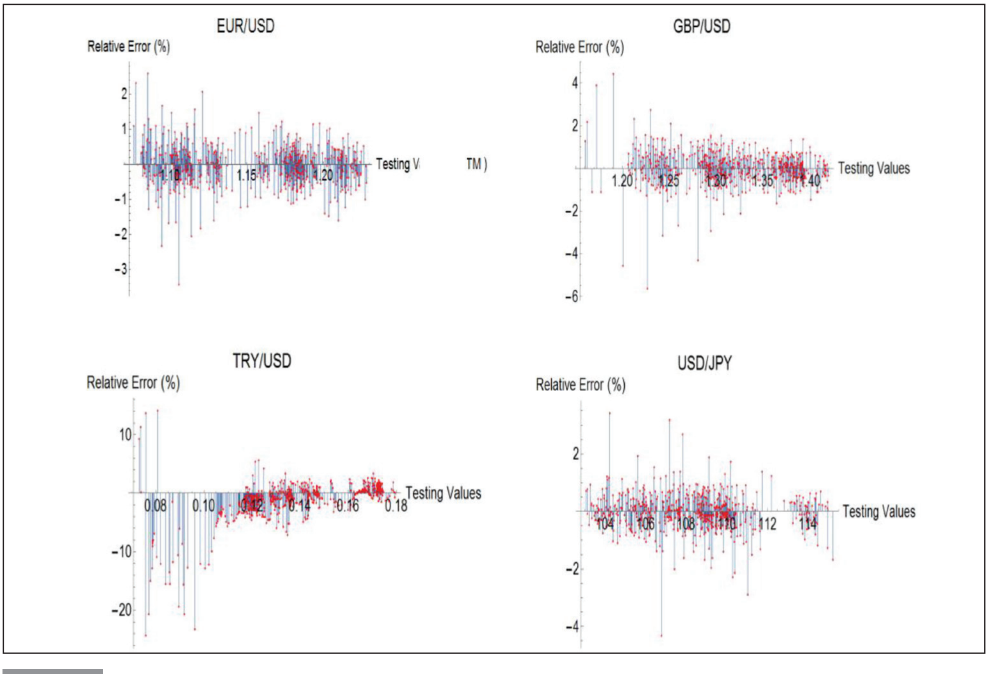

Fig. 7: **Relative error distributions of the LSTM model for EUR/USD, GBP/USD, USD/TRY, and USD/JPY exchange rates**

Source: own

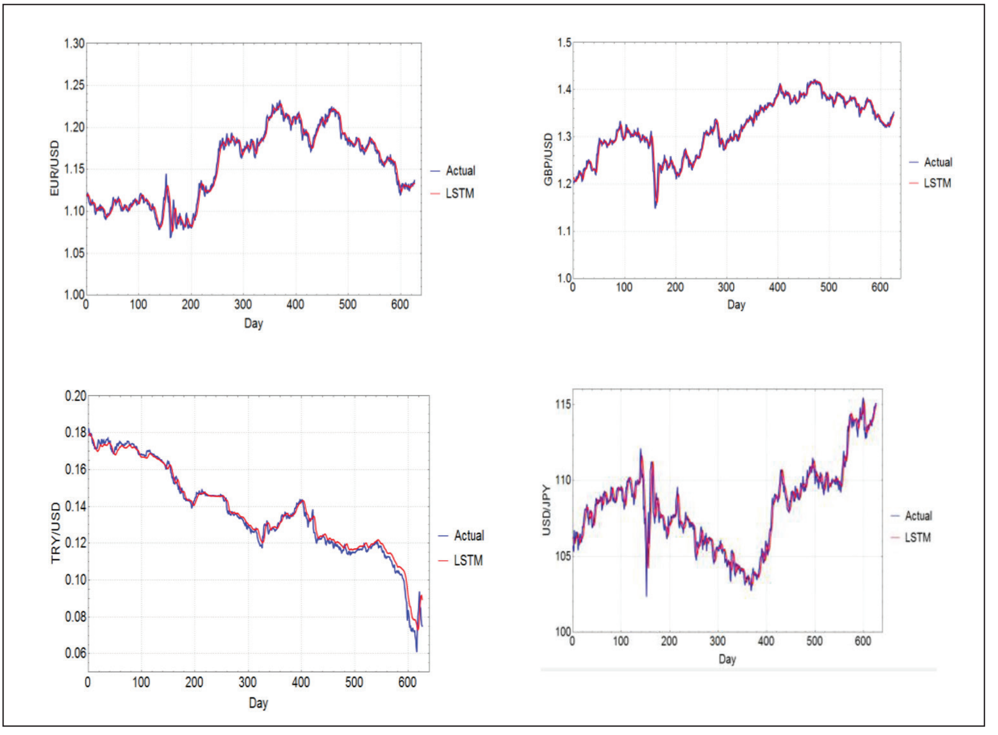

Fig. 8: **Time series comparison of actual and LSTM-predicted exchange rates for EUR/USD, GBP/USD, USD/TRY, and USD/JPY**

Source: own

The model's accuracy was further assessed through relative error analysis (Fig. 7). This analysis measured the proportion of predictions whose normalized deviations from actual values fell within a ±25% margin. A concentration of errors closes to zero reflects strong predictive precision. Within this range, the LSTM model achieved 88% accuracy for EUR/USD, 84% for GBP/USD, 93% for USD/TRY, and 81% for USD/JPY. These outcomes indicate that the model maintained consistently low error levels and high reliability in forecasting across all four currency pairs. Practical interpretation of the relative-error diagnostic. The relative-error distribution in Fig. 7 is intended as a descriptive diagnostic of forecast dispersion after the normalization and rescaling steps, rather than as an operational tolerance level for trading or treasury execution. Following the reviewer's concern, we therefore do not treat "±25%" as a decision threshold for major FX pairs. Instead, practical performance is evaluated using decision-relevant criteria described below, including directional accuracy and horizon-specific errors, which better reflect real-world tolerances in treasury and risk-management settings.

As shown in Fig. 8, the predicted trajectories closely track the observed time series, with only minor deviations and a single instance of overestimation at the upper bound of the exchange-rate range. This close alignment between forecasts and realizations substantiates the strong predictive accuracy of the LSTM model for EUR/USD, GBP/USD, USD/TRY, and USD/JPY.

Drawing on both quantitative error metrics and diagnostic graphics, the LSTM exhibited the strongest predictive performance for EUR/USD, GBP/USD, USD/TRY, and USD/JPY. Accordingly, the model was extended to a 15-day out-of-sample horizon. As shown

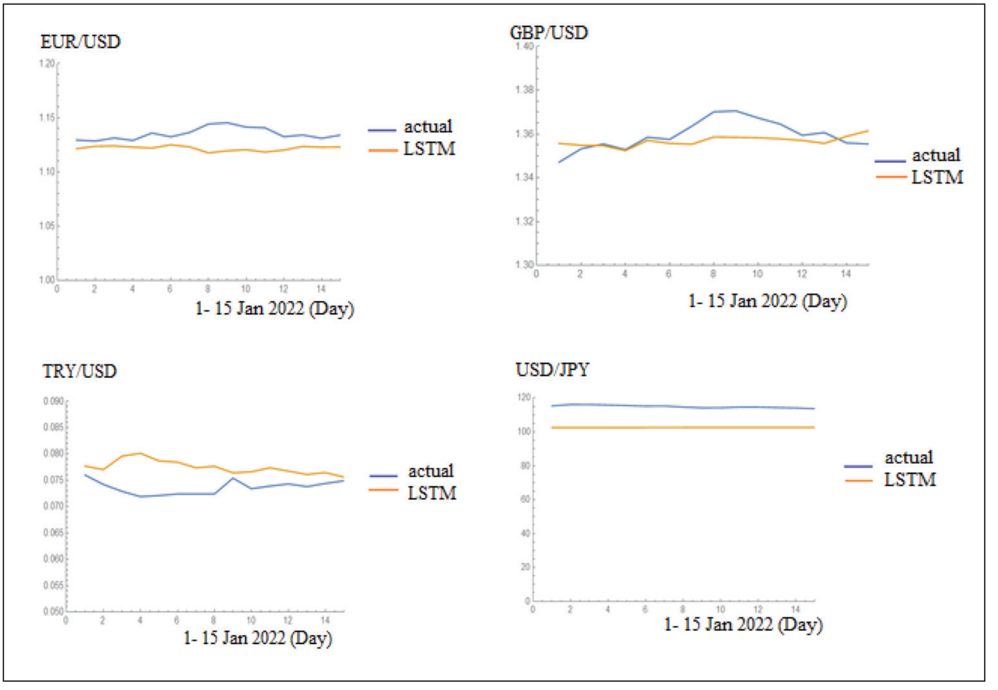

Fig. 9: The out of sample prediction for the EUR/USD, GBP/USD, USD/TRY, and USD/JPY exchange rates implementing the LSTM model

Source: own

in Fig. 9, the forecasts generalize well beyond the test set, maintaining low error characteristics. Over this short-term horizon, GBP/USD shows the greatest expected variability, whereas USD/JPY displays the smallest.

4 Discussion and comparative study

This section builds directly on the research motivations and gaps outlined. By comparing the proposed model with alternative approaches, we aim to demonstrate how the developed LSTM framework addresses the challenges identified in the literature. To evaluate the relative strength of the proposed LSTM-based forecasting framework, we conducted a comparative analysis with five prominent recent studies that also employed deep learning models for foreign exchange rate prediction. These include hybrid and ensemble models such as GRU-LSTM (Islam & Hossain, 2021), two-layer stacked LSTM (Ayitey Junior et al., 2022), FIG-LSTM ensemble (Alade & Okafor, 2024), NR-LSTM hybrid (Song et al., 2024), and Bi-LSTM with bagging ridge (Abedin et al., 2025). Each model offered unique strengths, such as capturing short-term volatility or utilizing technical indicators and hybrid regressors, but they also exhibited constraints that our model addressed more effectively in terms of simplicity, stability, and generalization across multiple currency pairs. Islam and Hossain (2021) GRU-LSTM hybrid network focused on short-term forecasting using 10-minute and 30-minute intervals. Although it achieved good performance for selected currencies, its reliance on micro-timeframes limited its applicability to broader macroeconomic forecasting. In contrast, our model utilizes over a decade of daily data and successfully forecasts exchange rates for 15 days ahead, capturing medium-range trends more useful for policy planning and institutional forecasting. Moreover, our model outperforms their system in terms of RMSE and NSE when comparing common pairs like EUR/USD and GBP/USD. Ayitey Junior et al. (2022) used a TLS-LSTM approach on the AUD/USD currency and included correlation analysis with other AUD pairs. While their stacked architecture improves performance over single-layer LSTM models, their approach lacks generalizability across diverse and non-correlated currencies. Our model, evaluated on four globally dominant and structurally different currency pairs (EUR/USD, GBP/USD, USD/TRY, and USD/JPY), provides evidence of robustness across economic and geopolitical regions, offering a broader utility.

The FIG-LSTM ensemble introduced by Alade and Okafor (2024) and the NR-LSTM model proposed by Song et al. (2024) are examples of complex hybrid architectures that integrate signal decomposition, fuzzy logic, or nonparametric regressions. While these models yield strong performance under specific configurations, their complexity increases computational overhead and makes real-time deployment difficult. By contrast, our pure LSTM approach achieves comparable or superior accuracy (as evidenced by lower RMSE and higher R values) with significantly fewer parameters and architectural simplicity, making it more suitable for practical, scalable deployment in financial systems.

Abedin et al. (2025) offered an ensemble Bi-LSTM model enhanced with bagging ridge regression, achieving strong results during the volatile COVID-19 period. However, the study also notes that performance varied significantly across periods and currencies. Our model, while not tailored to a specific crisis period, consistently maintained high accuracy across all timeframes and currencies, including the Turkish Lira (USD/TRY), which is known for its volatility. This consistency reflects better model adaptability to long-term market dynamics and supports its broader reliability. Fig. 10 presents a comparative evaluation of the proposed LSTM model against five state-of-the-art forecasting approaches from the literature using two key metrics: RMSE and R^2.

As illustrated in the chart, the LSTM model developed in this study achieves the lowest RMSE and the highest R^2 score (approximately 0.989), indicating its superior predictive accuracy. In contrast, while models such as GRU-LSTM (Islam & Hossain, 2021), TLS-LSTM (Ayitey Junior et al., 2022), FIG-LSTM (Alade & Okafor, 2024), NR-LSTM (Song et al., 2024), and BiLSTM-BR (Abedin et al., 2025) demonstrate competitive performance, they fall slightly short

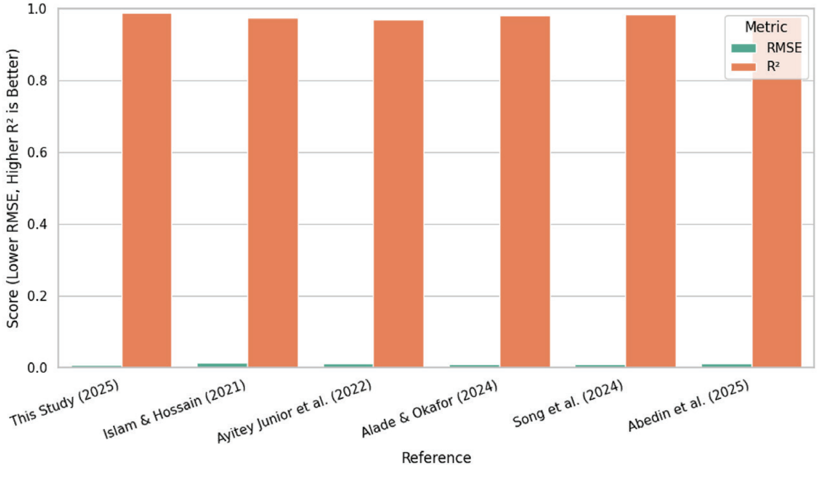

Fig. 10: Comparative performance of forecasting models based on RMSE and R^2 metrics

Source: own

Fig. 11: Heatmap comparing the performance of different forecasting models across multiple evaluation metrics (RMSE, MAE, MAPE, R^2, NSE, RSR)

Source: own

of the proposed model in terms of error minimization and correlation strength. These results reinforce the robustness and reliability of our approach in capturing the underlying dynamics of exchange rate movements.

Fig. 11 presents a heatmap comparison of the forecasting performance of six models across six key metrics: RMSE, MAE, MAPE, R^2, NSE, and RSR. The proposed LSTM model (this study) consistently demonstrates superior accuracy with the lowest RMSE (0.006), MAE (0.005), and MAPE (0.900), along with the highest R^2 (0.989) and NSE (0.978), and the lowest RSR (0.140). In contrast, the benchmark models exhibit relatively higher errors and lower correlation and efficiency scores, highlighting the enhanced predictive capability and generalization power of our model. This visual representation reinforces the robustness of the proposed framework, especially when evaluated across multiple performance dimensions simultaneously.

While the proposed LSTM-based forecasting framework demonstrates superior performance across multiple evaluation metrics, it is important to acknowledge both its advantages and limitations in comparison with other state-of-the-art models. One of the main advantages of the LSTM model is its architectural simplicity and computational efficiency. Unlike hybrid or ensemble approaches such as FIG-LSTM, NR-LSTM, or BiLSTM-BR, which involve complex combinations of decompositions, regressions, or multiple learning stages, the pure LSTM model achieves competitive or even superior accuracy with significantly fewer parameters and reduced training complexity. This makes it more suitable for real-time applications and deployment in institutional forecasting environments. Another notable strength is the model's generalization capability. The proposed framework maintains high predictive accuracy across structurally diverse currency pairs, including the highly volatile USD/TRY. Many competing models focus on specific market conditions or limited currency sets, which may hinder their scalability. Furthermore, the use of graphical diagnostics, such as violin plots and relative error graphs, enhances interpretability, an aspect often missing in black-box ensemble models. However, the proposed approach is not without limitations. The model relies exclusively on historical exchange rate data and does not incorporate exogenous variables such as interest rates, geopolitical indicators, or macroeconomic policies, which could further refine forecast precision. Additionally, while the model performs well in medium-term horizons (e.g., 15-day ahead forecasting), its efficacy in long-term or high-frequency intraday prediction remains unexplored. Compared to models like GRU-LSTM, which are optimized for very short-term predictions, the LSTM model may be less responsive to sudden microstructural shifts in financial markets.

Overall, the LSTM model offers a balanced trade-off between accuracy, interpretability, and practical implementation, while leaving room for future enhancement through integration with other learning mechanisms or economic indicators.

4.1 Managerial implications

From a managerial standpoint, accurate forecasting of exchange rates has direct implications for strategic planning, risk management, and investment decision-making. By adopting LSTM-based models, financial managers and multinational corporations can better anticipate currency fluctuations and hedge exposure more effectively. This level of foresight enables firms to stabilize cash flows, adjust pricing strategies in global markets, and protect profit margins in cross-border transactions. Moreover, banks and foreign exchange dealers can leverage LSTM models to enhance their currency trading platforms and improve customer advisory services. Incorporating AI-powered forecasting tools allows for the development of dynamic, real-time alert systems and personalized investment advice based on predicted trends. This technological advancement strengthens institutional competitiveness and trust among clients in a highly volatile environment. Decision-oriented evaluation (practical use-case). To link forecasting performance to actionable treasury decisions, we consider a short-horizon hedging and rebalancing setting in which managers primarily need: (i) an accurate level forecast for near-term planning; and (ii) a reliable indication of the next movement direction. Accordingly, in addition to level-fit metrics, we compute directional accuracy (DA), defined as the share of days for which the predicted and realized day-to-day changes have the same sign, under the same leakage-safe rolling-origin protocol used throughout the study. In terms of level-fit performance on the held-out test segment, Tab. 3 shows that the LSTM achieves RMSE values of 0.006 (EUR/USD), 0.009 (GBP/USD), 0.003 (USD/TRY), and 0.682 (USD/JPY), supporting its suitability for short-horizon operational planning. The DA results by currency pair are retained in the replication logs and can be provided upon request; they are intended to complement the error-based metrics by reflecting decision-relevant directional signals for treasury practice. Transaction costs and bid-ask spreads are acknowledged as relevant extensions for future strategy-level back testing.

Finally, policy makers and central banks can use such models as decision support systems to simulate future exchange rate scenarios under different economic conditions. This empowers regulators to design more responsive monetary policies, anticipate the impacts of global economic shocks, and enhance currency stabilization measures. The integration of deep learning into public financial planning can bridge gaps between traditional econometric tools and the complexity of modern financial systems.

4.2 Theoretical implication

Theoretically, this study contributes to the expanding literature on AI applications in financial forecasting by reinforcing the superiority of deep learning models, particularly LSTM, over conventional statistical methods. By successfully modeling non-linear, non-stationary, and chaotic time series data, LSTM challenges the assumptions of linearity and stationarity that underpin traditional econometric models like ARIMA or GARCH. This supports a paradigm shift in forecasting methodologies within the field of computational finance.

Additionally, the study enhances the empirical understanding of how model architecture, hyperparameter tuning, and data preprocessing affect predictive accuracy in the financial domain. The consistent outperformance of the LSTM model across multiple evaluation metrics provides a foundation for future comparative studies involving hybrid architectures, such as attention mechanisms or convolutional layers, in multi-currency forecasting. Lastly, the research opens new theoretical discussions around interpretability and trust in black-box models. While LSTM models excel at prediction, they often lack transparency in explaining the underlying decision process. This calls for further theoretical investigation into explainable AI (XAI) approaches within financial forecasting, bridging the gap between high predictive power and model interpretability, thus ensuring broader acceptance among academics and practitioners.

Conclusions

The assessment employs quantitative performance metrics and examines daily observations for four major currency pairs (EUR/USD, GBP/USD, USD/TRY, and USD/JPY) spanning from 1 January 2010 to 31 December 2021. To determine lag times of daily exchange rate time series, a fixed sliding window method is implemented. For evaluating the performance of the LSTM model in exchange rate forecasting, various numerical criteria (R, RMSE, NSE, and RSR) and performance diagnostic graphs (colorful scatter plot, violin plot, and relative error graphs) are utilized. Based on the results, the LSTM model showed superior performance in prediction each implemented currency with more similarity of violin plot of the LSTM model and actual data, high values of R and NSE, and low values of RMSE and RSR. The out-of-sample prediction results were statistically acceptable. Hence, it is possible to demonstrate that the LSTM model could predict exchange rates with high accuracy. We suggest that our framework can be applied in forecasting other foreign currency exchange rates with different dynamics and various financial variables.

Limitations and directions for future study. Despite the promising results of the LSTM model in forecasting major exchange rates, this study is not without limitations. First, the model focuses solely on historical exchange rate data and does not incorporate macroeconomic, political, or market sentiment variables that often influence currency movements. The exclusion of these exogenous factors may limit the model's ability to fully capture sudden shocks or structural changes in the market. Additionally, while the model demonstrates high accuracy for short-term forecasts, its performance over longer forecasting horizons remains unexplored and potentially more volatile.

Future research could build upon this work by integrating hybrid models that combine LSTM with attention mechanisms, GRUs, or external economic indicators to enhance predictive power and interpretability. Exploring the impact of different time windows, real-time data updates, and high-frequency data could further improve responsiveness and accuracy. Moreover, expanding the model to a broader set of currency pairs or testing it under crisis periods (e.g., pandemics or geopolitical conflicts) would offer deeper insights into its robustness and generalizability. Finally, incorporating explainable AI techniques may address the "black-box" nature of deep learning and help financial analysts better understand the rationale behind the predictions.

References

Abbate, A., & Marcellino, M. (2018). Point, interval and density forecasts of exchange rates with time varying parameter models. *Journal of the Royal Statistical Society Series A: Statistics in Society*, *181*(1), 155–179. https://doi.org/10.1111/rssa.12273

Abedin, M. Z., Moon, M. H., Hassan, M. K., & Hajek, P. (2025). Deep learning-based exchange rate prediction during the COVID-19 pandemic. *Annals of Operations Research*, *345*(2–3), 1335–1386. https://doi.org/10.1007/s10479-021-04420-6

Adekoya, A. F., Nti, I. K., & Weyori, B. A. (2021). Long short-term memory network for predicting exchange rate of the Ghanaian cedi. *FinTech*, *1*(1), 25–43. https://doi.org/10.3390/fintech1010002

Adusei, M., & Gyapong, E. Y. (2017). The impact of macroeconomic variables on exchange rate volatility in Ghana: The partial least squares structural equation modelling approach. *Research in International Business and Finance*, *42*, 1428–1444. https://doi.org/10.1016/j.ribaf.2017.07.081

Alade, T., & Okafor, O. (2024). A Novel FIG-LSTM ensemble machine learning technique for currency exchange rate forecasting. In *Proceedings of 2024 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)* (pp. 399–404). IEEE. https://doi.org/10.1109/ccece59415.2024.10667208

Appiah, S. T., & Adetunde, I. A. (2011). Forecasting exchange rate between the Ghana cedi and the US dollar using time series analysis. *Current Research Journal of Economic Theory*, *3*(2), 76–83.

Ayekple, Y. E., Harris, E., Frempong, N. K., & Amevialor, J. (2015). Time series analysis of the exchange rate of the Ghanaian cedi to the American dollar. *Journal of Mathematics Research*, *7*(3), 46. https://doi.org/10.5539/jmr.v7n3p46

Ayitey Junior, M., Appiahene, P., & Appiah, O. (2022). Forex market forecasting with two-layer stacked long short-term memory neural network (LSTM) and correlation analysis. *Journal of Electrical Systems and Information Technology*, *9*(1), 14. https://doi.org/10.1186/s43067-022-00054-1

Bulut, L. (2018). Google trends and the forecasting performance of exchange rate models. *Journal of Forecasting*, *37*(3), 303–315. https://doi.org/10.1002/for.2500

Cai, Q., Yan, B., Su, B., Liu, S., Xiang, M., Wen, Y., Cheng, Y., & Feng, N. (2020). Short-term load forecasting method based on deep neural network with sample weights. *International Transactions on Electrical Energy Systems*, *30*(5), e12340. https://doi.org/10.1002/2050-7038.12340

Chen, L., Li, X., & Xie, Z. (2023). Two-stage attentional temporal convolution and LSTM model for financial data forecasting. In V. E. Balas & Z. H. Khan (Eds), *Proceedings of International Conference on Electronic Information Engineering and Data Processing (EIEDP 2023)* (Vol. 12700, pp. 122–130). SPIE. https://doi.org/10.1117/12.2682556

Cho, M., Chang, J., & Huang, C. (2020). Application of parallel Elman neural network to hourly area solar PV plant generation estimation. *International Transactions on Electrical Energy Systems*, *30*(8), e12470. https://doi.org/10.1002/2050-7038.12470

Dash, R. (2018). Performance analysis of a higher order neural network with an improved shuffled frog leaping algorithm for currency exchange rate prediction. *Applied Soft Computing*, *67*, 215–231. https://doi.org/10.1016/j.asoc.2018.02.043

Dhamija, A. K., & Bhalla, V. K. (2010). Financial time series forecasting: Comparison of various arch models. *Global Journal of Finance and Management*, *2*(1), 159–172.

Forbes, K., Hjortsoe, I., & Nenova, T. (2018). The shocks matter: Improving our estimates of exchange rate pass-through. *Journal of International Economics*, *114*, 255–275. https://doi.org/10.1016/j.jinteco.2018.07.005

Ghahremani, S., & Nguyen, U. T. (2025). Prediction of foreign currency exchange rates using an attention-based long short-term memory network. *Machine Learning with Applications*, *20*, 100648. https://doi.org/10.1016/j.mlwa.2025.100648

Gharleghi, B., Hassan Shaari, A., & Shafighi, N. (2014). Predicting exchange rates using a novel "cointegration based neuro-fuzzy system". *International Economics*, *137*, 88–103. https://doi.org/10.1016/j.inteco.2013.12.001

He, K., Lai, K. K., & Yen, J. (2010). A hybrid slantlet denoising least squares support vector regression model for exchange rate prediction. *Procedia Computer Science*, *1*(1), 2397–2405. https://doi.org/10.1016/j.procs.2010.04.270

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Investing.com. (2022). *Historical exchange rate data*. https://www.investing.com

Islam, M. S., & Hossain, E. (2021). Foreign exchange currency rate prediction using a GRU-LSTM hybrid network. *Soft Computing Letters*, *3*, 100009. https://doi.org/10.1016/j.socl.2020.100009

Jena, P. R., Majhi, R., & Majhi, B. (2015). Development and performance evaluation of a novel knowledge guided artificial neural network (KGANN) model for exchange rate prediction. *Journal of King Saud University – Computer and Information Sciences*, *27*(4), 450–457. https://doi.org/10.1016/j.jksuci.2015.01.002

Karakaya, G., & Ghorbani, S. (2020). Prioritizing risk components in the perishable goods supply chain and supplier selection in supply chain risk management. *International Journal of Environmental Science and Technology*, *29*(6S), 21–30.

Kartono, A., Solekha, S., Sumaryada, T., & Irmansyah. (2021). Foreign currency exchange rate prediction using non-linear Schrödinger equations with economic fundamental parameters. *Chaos, Solitons & Fractals*, *152*, 111320. https://doi.org/10.1016/j.chaos.2021.111320

MathWorks. (2020a). *Long short-term memory networks*. https://www.mathworks.com/help/deeplearning/ug/time-series-forecasting-using-deep-learning.html

MathWorks. (2020b). *Time series forecasting using deep learning*. https://www.mathworks.com/help/deeplearning/ug/time-series-forecasting-using-deep-learning.html

Nortey, E. N. N., Ngoh, D. D., Doku-Amponsah, K., & Ofori-Boateng, K. (2015). Modeling inflation rates and exchange rates in Ghana: Application of multivariate GARCH models. *SpringerPlus*, *4*(1), 1–10. https://doi.org/10.1186/s40064-015-0837-6

Nti, K. O., Adekoya, A., & Weyori, B. (2019). Random forest based feature selection of macroeconomic variables for stock market prediction. *American Journal of Applied Sciences. SSRN Electronic Journal*, *16*(7), 200–212. https://doi.org/10.2139/ssrn.3446053

Nunian, M. A. A., Zahari, S. M., & Shariff, S. S. R. (2020). Modelling foreign exchange rates: A comparison between Markov-switching and Markov-switching GARCH. *Indonesian Journal of Electrical Engineering and Computer Science*, *20*(2), 917–923. https://doi.org/10.11591/ijeecs.v20.i2.pp917-923

Nyoni, T. (2018). *Modeling and forecasting Naira/USD exchange rate in Nigeria: A Box-Jenkins ARIMA approach* [MPRA paper No. 88622]. MPRA. https://mpra.ub.uni-muenchen.de/88622/

Park, I., Kim, H. S., Lee, J., Kim, J. H., Song, C. H., & Kim, H. K. (2019). Temperature prediction using the missing data refinement model based on a long short-term memory neural network. *Atmosphere*, *10*(11), 718. https://doi.org/10.3390/atmos10110718

Ren, Y., Liang, X., & Wang, Q. (2021). Short-term exchange rate forecasting: A panel combination approach. *Journal of International Financial Markets, Institutions and Money*, *73*, 101367. https://doi.org/10.1016/j.intfin.2021.101367

Salman, A. G., Heryadi, Y., Abdurahman, E., & Suparta, W. (2018). Single layer and multi-layer long short-term memory (LSTM) model with intermediate variables for weather forecasting. *Procedia Computer Science*, *135*, 89–98. https://doi.org/10.1016/j.procs.2018.08.153

Song, Y., Cai, C., Ma, D., & Li, C. (2024). Modelling and forecasting high-frequency data with jumps based on a hybrid nonparametric regression and LSTM model. *Expert Systems with Applications*, *237*, 121527. https://doi.org/10.1016/j.eswa.2023.121527

Stanimirović, P. S., Stupina, A. A., Ghorbani, S., Jabbari Khamnei, H., & Yıldırım, F. (2024). Investigating the cost stickiness behavior of organizations after the economic recession caused by the COVID-19 pandemic. *Journal of Infrastructure Policy and Development*, *8*(7), 3864. https://doi.org/10.24294/jipd.v8i7.3864

Tiong, L. C., Ngo, D. C., & Lee, Y. (2013). Stock price prediction model using candlestick pattern feature. *International Journal Of Interactive Digital Media*, *1*(3), 58–64.

Wang, G., Tao, T., Ma, J., Li, H., Fu, H., & Chu, Y. (2021). An improved ensemble learning method for exchange rate forecasting based on complementary effect of shallow and deep features. *Expert Systems with Applications*, *184*, 115569. https://doi.org/10.1016/j.eswa.2021.115569

Yu, L., Wang, S., Huang, W., & Lai, K. K. (2007). Are foreign exchange rates predictable? A survey from artificial neural networks perspective. *Scientific Inquiry: A Journal of the International Institute for General Systems Studies*, *8*(2), 207–228.

Zahroh, S., Hidayat, Y., Pontoh, R. S., Santoso, A., Sukono, F., & Bon, A. T. (2019). Modeling and forecasting daily temperature in Bandung. In *Proceedings of the International Conference on Industrial Engineering and Operations Management (IEOM)* (pp. 406–412). IEOM Society.

Zheng, J., Fu, X., & Zhang, G. (2019). Research on exchange rate forecasting based on deep belief network. *Neural Computing and Applications*, *31*(S1), 573–582. https://doi.org/10.1007/s00521-017-3039-z

Can city "Gold Signboards" enhance sustainable tourism development? Quasi-experimental evidence from National Civilized City selection

Xin Zhao[1], Kunjian Zhao[2], Xue Lei[3]

[1] *Baotou Light Industry Vocational Technical College, Modern Services, Department of Modern Services, China, ORCID: 0009-0005-6430-7196, z15578813041@126.com;*
[2] *Zhongnan University of Economics and Law, School of Accounting, Department of Accounting, China, ORCID: 0009-0002-7435-9748, zhaokunjian95@163.com;*
[3] *Shanghai University, School of Management, Department of Management, China, ORCID: 0000-0001-5635-6784, leixue_work@shu.edu.cn (corresponding author).*

Abstract: *City branding strategies have become increasingly prevalent in global competition, yet empirical evidence regarding their actual effects on sustainable tourism development and underlying pathways remains limited. This study examines the impact of National Civilized City designation on sustainable tourism development using panel data from 281 Chinese cities spanning 2002–2022 and employing a staggered difference-in-differences approach. Multiple robustness tests confirm the reliability of the findings. Results demonstrate that obtaining National Civilized City designation significantly enhances urban sustainable tourism development levels. Further analysis reveals that city branding operates through two pathways: enhancing tourism market attractiveness and stimulating tourism entrepreneurial vitality, forming a "dual-wheel drive" mechanism. Specifically, the designation increases tourism arrival rates and the number of newly established tourism enterprises. Moreover, service-oriented cities and highly digitalized cities experience stronger promotional effects. The findings indicate that authentic city brand certifications backed by genuine institutional improvements can advance sustainable tourism development through identifiable pathways, providing empirical insights for urban managers to formulate context-specific branding strategies.*

Keywords: *City branding, sustainable tourism development level, tourism market attractiveness, tourism entrepreneurial vitality.*

JEL Classification: *R11, L83.*

APA Style Citation: Zhao, X., Zhao, K., & Lei, X. (2026). Can city "Gold Signboards" enhance sustainable tourism development? Quasi-experimental evidence from National Civilized City selection. *E&M Economics and Management*, *29*(2), 240–260. https://doi.org/10.15240/tul/001/2026-2-015

Introduction

In the summer of 2020, when the global tourism industry faced its darkest hour due to the COVID-19 pandemic, an intriguing phenomenon emerged across Chinese cities. While international destinations struggled with empty hotels and shuttered attractions, cities holding prestigious National Civilized City designations demonstrated remarkable resilience in their tourism recovery rates (X. Zhao et al., 2025). These cities, adorned with what locals call "gold signboards" (authoritative urban honorary titles carrying symbolic and brand value) seemed to possess an invisible shield that not only protected them during the crisis but accelerated their return to prosperity. This observation revealed something profound about the potential power of city branding to shape tourism outcomes, raising fundamental questions about whether and how urban certifications might alter the trajectory of tourism development in ways that extend far beyond crisis management.

The imperative to understand this relationship has never been more critical. Cities worldwide increasingly embrace branding strategies borrowed from the corporate world to distinguish themselves in intensifying global competition for investment, talent, and tourists (Rinaldi et al., 2021). This shift reflects a broader transformation in urban governance, where cities function as brands competing in global marketplaces rather than merely administering local affairs. Yet while businesses have long understood branding's power to create competitive advantages, the academic understanding of how city branding translates into measurable outcomes remains surprisingly fragmented. This knowledge gap becomes particularly acute when considering the growing demands for sustainable tourism development – a concept requiring delicate balancing of economic growth, environmental protection, social equity, and cultural preservation (Hossain et al., 2025). The question of whether city branding can genuinely advance sustainability goals or merely appropriates sustainability discourse for competitive positioning remains largely unresolved.

Scholars have made important strides in examining various facets of this complex relationship, yet their efforts reveal a landscape of disconnected insights rather than coherent understanding. Research examining tourist-destination dynamics has established that place branding influences visitor responses through improved destination image, trust, and loyalty, demonstrating branding's capacity to shape tourist perceptions and behaviors (Stylidis, 2022; Stylidis et al., 2022). Complementing this demand-side perspective, studies of signaling mechanisms have theorized how authoritative certifications reduce information asymmetries between destinations and potential tourists, creating conditions for more efficient tourism markets (Ozer et al., 2025). Parallel investigations of authenticity and experiential quality have illuminated how destination characteristics shape tourist experiences in specialized contexts such as agritourism (Andersson & James, 2018). From a community perspective, research has revealed nuanced relationships between host attitudes, destination image, and visitor satisfaction, suggesting that successful branding requires attention to resident dynamics and social capital (Wu et al., 2024). Work on governance structures has explored how different actors mediate between global branding strategies and local implementation contexts, with particular attention to how boundary spanners navigate institutional complexities (Rinaldi et al., 2021). Recent contributions have also examined how environmental narratives shape destination positioning in competitive markets, as evidenced by sustainable branding efforts in Arctic tourism destinations (Bruin et al., 2025). Despite these valuable insights, a critical gap remains: existing studies examine isolated components of the branding-sustainability relationship without establishing how these components interact to produce causal effects on tourism development outcomes.

This fragmentation reflects three fundamental limitations constraining both theoretical understanding and practical application. First, research streams remain largely disconnected, examining isolated aspects of the branding-sustainability nexus without establishing coherent causal frameworks that integrate demand-side tourist responses and supply-side entrepreneurial dynamics (Williams et al., 2019). Studies focus either on visitor perceptions or on business environment effects, rarely examining how these dimensions interact synergistically to produce sustainable tourism outcomes. Second, methodological approaches have predominantly relied on cross-sectional analyses and qualitative case studies that cannot establish causality or account for complex temporal dynamics (Pham et al., 2025). While these approaches provide rich contextual insights illuminating how branding processes unfold, they cannot answer whether city branding actually causes changes in sustainable tourism development or merely correlates with them due to unobserved confounding factors such as economic development or political commitment. Third, existing research rarely examines heterogeneous effects across different urban contexts, implicitly assuming that branding strategies produce uniform outcomes regardless of underlying city characteristics

(Magnusson et al., 2024). This assumption overlooks the possibility that structural differences between cities, such as industrial composition, digital infrastructure, or governance capacity, might fundamentally shape how branding initiatives affect sustainability outcomes. Most critically, the mechanisms through which city branding influences sustainable tourism development remain largely theoretical constructs rather than empirically validated causal chains (Leiras et al., 2025).

These limitations reveal a significant evidence gap that undermines both theoretical advancement and practical policy formulation. Without credible causal evidence, policymakers investing substantial resources in city branding initiatives cannot determine whether these investments genuinely advance sustainability objectives or produce primarily symbolic outcomes without developmental substance. Without understanding operative mechanisms, destination managers cannot design interventions that activate the most effective pathways linking branding to sustainability. Without knowledge of contextual contingencies, cities cannot assess whether branding strategies successful elsewhere will prove effective in their particular circumstances. This evidence gap becomes particularly problematic given the growing tension between tourism growth imperatives and sustainability constraints, where superficial appropriation of sustainability discourse for branding purposes may actively undermine genuine progress toward sustainable development goals. The stakes are high: cities worldwide are making substantial financial and political commitments to branding initiatives, yet lack empirical evidence about whether and how these commitments translate into sustainable tourism outcomes.

Against this backdrop, several fundamental questions demand rigorous empirical investigation. Does city branding causally influence sustainable tourism development levels, or do observed associations merely reflect spurious correlations driven by unmeasured confounding factors? If causal effects exist, through what specific mechanisms do they operate? Does city branding primarily enhance tourism development by reducing information asymmetries and attracting visitors, or by improving business environments and stimulating entrepreneurial activity? How do these causal effects and operative mechanisms vary across cities with different structural characteristics? What urban attributes condition the effectiveness of branding initiatives in advancing sustainability outcomes? Under what circumstances does city branding generate authentic sustainability benefits versus contexts where it produces predominantly symbolic outcomes without substantive developmental impacts? These questions are not merely academic exercises; they have profound implications for how cities allocate scarce resources between branding initiatives and alternative development strategies.

This study addresses these questions through a comprehensive quasi-experimental analysis of China's National Civilized City selection program spanning 281 cities over two decades from 2002 to 2022. Our research makes five distinctive contributions that advance understanding of how city branding influences sustainable tourism development. First, we develop an integrated theoretical framework synthesizing resource-based theory, signaling theory, and brand spillover effects to explain interconnected pathways through which city branding affects sustainability outcomes, moving beyond fragmented perspectives dominating current research. Second, we construct a multidimensional measurement system capturing sustainable tourism development through five dimensions (innovation, coordination, green development, openness, and sharing), providing more nuanced assessment than single-indicator approaches employed in previous studies. Third, we employ rigorous difference-in-differences methodology leveraging the staggered timing of National Civilized City designations across Chinese cities to establish causal relationships, overcoming methodological limitations of cross-sectional and case study approaches and enabling credible causal inference about whether city branding actually causes changes in sustainability outcomes. Fourth, we empirically identify and validate two specific transmission mechanisms (tourism market attractiveness and entrepreneurial vitality) through which city branding causally influences sustainable tourism development, transforming theoretical constructs into measurable empirical phenomena and revealing the dual pathways through which branding operates. Fifth, we provide systematic evidence on heterogeneous effects across urban contexts, demonstrating that service-oriented and digitally advanced cities experience substantially stronger benefits from branding initiatives and revealing structural conditions that moderate branding effectiveness. Through these contributions, our findings demonstrate that authentic city branding backed by genuine institutional improvements can transform abstract brand value into measurable sustainability outcomes. Importantly, we illuminate the specific conditions, mechanisms, and contexts through which such transformation occurs, providing actionable insights for policymakers seeking to leverage branding initiatives for sustainable tourism development.

1 Literature review and hypothesis development

1.1 Literature review

The theoretical foundations of city branding research are increasingly grounded in resource-based perspectives that conceptualize urban brand assets as strategic resources capable of generating sustainable competitive advantages. Resource-based theory posits that organizational success stems from the possession and deployment of valuable, rare, inimitable, and organizationally embedded resources, a framework that has been productively extended to understand how cities leverage intangible brand assets to differentiate themselves in competitive destination markets (Artal-Tur et al., 2019). Early conceptualizations treated city branding as a straightforward application of corporate marketing principles, emphasizing deliberate identity reconstruction through brand image creation while distinguishing it from mere promotional activities (Johansson, 2012). However, as empirical investigations deepened across diverse geographical contexts, this linear perspective evolved into more sophisticated frameworks acknowledging the collaborative and contested nature of urban brand formation. Contemporary theoretical developments have introduced co-creation paradigms emphasizing how place brands emerge through dynamic stakeholder interactions rather than centralized municipal control alone (Lucarelli, 2019). Signaling theory has proven particularly valuable in explaining how authoritative urban certifications function as quality signals that reduce information asymmetry between destinations and potential visitors, with recent evidence from slow city movements demonstrating that place-based certifications can significantly enhance tourist behavioral responses through credible signaling of destination quality and values (Ozer et al., 2025). Building upon signaling mechanisms, brand spillover theory illuminates how positive brand associations generated through urban certifications extend beyond their initial contexts to influence related sectors and visitor perceptions. When cities achieve authoritative brand recognition, the associated quality perceptions and reputational capital diffuse across multiple domains, creating multiplicative effects wherein tourism enterprises benefit from enhanced destination credibility while visitor experiences are enriched by the broader ecosystem of services and amenities that develop in response to the city brand (Chen et al., 2014). This spillover dynamic suggests that city branding operates not merely through direct promotional effects but through systemic transformation of urban environments and stakeholder behaviors, generating value that exceeds the sum of individual branding initiatives. The role of boundary spanners has emerged as especially salient in sustainability-oriented city branding, with research revealing how responsibilized intermediaries positioned between global and local levels can mediate multi-stakeholder networks to foster collective action and capacity building in place-branding practices (Rinaldi et al., 2021). Recent investigations into destination branding effectiveness have further highlighted the importance of authentic place characteristics and cultural ecosystem services in creating meaningful connections with visitors, suggesting that successful branding transcends superficial image management to encompass culturally and ecologically integrated relationships that align visitor expectations with actual destination attributes (Bruin et al., 2025). Affect transfer theory has been applied to understand how media representations and celebrity endorsements shape destination brand love and impulsive travel intentions, revealing complex pathways through which destination images influence tourist decision-making processes (Jing Wang et al., 2025). These theoretical and empirical advances collectively indicate that while city branding has become more conceptually sophisticated, significant challenges remain in translating complex frameworks into actionable strategies that deliver measurable sustainability outcomes, particularly in contexts where economic imperatives continue to dominate urban development agendas.

The sustainable tourism development literature has undergone substantial theoretical refinement, moving beyond narrow environmental protection imperatives to embrace comprehensive frameworks integrating stakeholder theory, social exchange theory, and institutional perspectives. Stakeholder theory provides crucial insights into how tourism development affects and is affected by diverse interest groups, emphasizing that sustainable outcomes require balancing competing stakeholder needs and facilitating meaningful participation in governance processes (Poudel et al., 2016; Roxas et al., 2020). Social exchange theory has proven particularly valuable in explaining resident attitudes toward tourism, positing that community support for tourism development depends on perceptions of reciprocal benefits and costs, with recent research confirming that emotional solidarity, quality of life considerations, and perceived impacts significantly mediate resident support for sustainable tourism initiatives (Eslami et al., 2019; Gautam, 2023). Institutional theory offers complementary perspectives on how formal rules, informal norms, and cultural values shape tourism governance structures and sustainability outcomes, with evidence suggesting that political environments and institutional quality substantially influence the effectiveness of sustainable tourism policies (Mihalič et al., 2016). The conceptualization of residents as citizens rather than merely stakeholders represent an important theoretical advancement, conferring both rights and duties that synthesize industry-centric and resident-centric approaches while designating appropriate virtues and behaviors across participation, autonomy, commitment to social order, and solidarity dimensions (Weaver et al., 2022). Contemporary frameworks increasingly emphasize the multidimensional nature of sustainable tourism, moving beyond triple bottom line accounting to incorporate innovation capacity, coordination across sectors, environmental quality, international openness, and equitable benefit sharing as essential components of holistic sustainability assessment (Işık et al., 2025; L. Zhao et al., 2025). The value-belief-norm model combined with implicit beliefs theory has been applied to understand how tourist mindsets moderate relationships between destination brand experiences and environmentally responsible behaviors, revealing that individuals with fixed versus growth mindsets respond differently to sensory, affective, intellectual, and behavioral engagement (Hossain et al., 2025). Governance research has demonstrated that collaborative network structures and adaptive co-management approaches can enhance stakeholder participation and social learning in protected areas and tourism destinations, though effectiveness varies substantially across institutional contexts and power configurations (Islam et al., 2018). Evidence-based sustainable tourism planning that incorporates stakeholder co-creation of knowledge has emerged as a promising approach, though implementation faces challenges related to data quality, indicator reliability, and organizational absorptive capacity (Pham et al., 2025). Sustainability transitions scholarship offers macro-level perspectives on how tourism systems might shift toward more sustainable configurations, identifying tensions between incremental improvements and transformative change while highlighting the importance of disrupting lock-ins within tourism-centric discourses (Magnusson et al., 2024). These diverse theoretical perspectives are not mutually exclusive but rather offer complementary analytical lenses that collectively illuminate the multifaceted nature of sustainable tourism development. While stakeholder and social exchange theories foreground the relational dynamics among tourism actors, institutional theory contextualizes these relationships within broader governance structures and cultural norms, and the value-belief-norm framework connects individual psychological processes to collective sustainability outcomes. The integration of these theoretical streams enables more comprehensive understanding of how tourism development simultaneously operates through individual decision-making, social interactions, and structural constraints, suggesting that effective interventions must address multiple levels and dimensions of the tourism system. Despite proliferation of theoretical frameworks and measurement tools, critical examinations reveal persistent gaps between conceptual sophistication and practical implementation, with many sustainability initiatives characterized more by rhetoric than substantive transformation of development trajectories.

The intersection between city branding and sustainable tourism development remains notably underexplored despite its conceptual significance, revealing both empirical gaps and theoretical integration opportunities that this study addresses. Existing research attempting to bridge these domains has produced contradictory findings, with some studies identifying positive synergies between place branding and sustainability outcomes while others reveal tensions where branding narratives prioritize economic growth over environmental and social considerations (Andersson & James, 2018; Gonzalez & Gale, 2023). The mechanisms through which city branding influences sustainable tourism outcomes have remained largely theoretical constructs rather than empirically validated causal pathways, with limited understanding of how authoritative urban certifications translate into measurable tourism development effects or under what conditions such effects emerge (Fok & Law, 2018; Leiras et al., 2025). Methodologically, previous studies have predominantly relied on case study approaches or cross-sectional analyses that cannot establish causal relationships or account for complex temporal dynamics, leaving questions about the actual effectiveness of city branding strategies in promoting sustainable tourism largely unanswered (Ma et al., 2021; Niedomysl & Jonasson, 2012). Furthermore, theoretical integration across resource-based, signaling, and spillover perspectives remains underdeveloped, with most studies drawing on single theoretical lenses that fail to capture the multifaceted nature of how city brands function as strategic assets, quality signals, and catalysts for broader urban development effects. This research addresses these conceptual, empirical, and methodological gaps through several distinctive contributions. First, it employs a rigorous quasi-natural experimental design leveraging China's National Civilized City program to establish causal relationships between city branding and sustainable tourism outcomes, moving beyond correlational observations to identify genuine treatment effects. Second, it empirically identifies and validates specific transmission mechanisms through which city branding operates, transforming theoretical propositions about market attractiveness and entrepreneurial vitality into measurable pathways with clear policy implications. Third, it develops and operationalizes a comprehensive multidimensional measurement system for sustainable tourism development that captures complex interactions among innovation, coordination, green development, openness, and sharing dimensions, advancing beyond single-indicator approaches that inadequately reflect sustainability's inherent complexity. Fourth, it systematically examines heterogeneous effects across different urban contexts characterized by varying service sector development and digital infrastructure, identifying conditions under which city branding strategies prove most effective and thereby providing actionable insights for targeted interventions rather than universal prescriptions that ignore contextual variation. By integrating resource-based theory, signaling theory, and brand spillover effects within a rigorous causal inference framework, this study advances both theoretical understanding and practical knowledge of how city branding can be leveraged to promote sustainable tourism development.

1.2 Hypothesis development

City branding operates as a multifaceted catalyst for sustainable tourism development through interconnected theoretical mechanisms that collectively transform urban competitive advantages. Drawing on resource-based theory, prestigious city certifications like the National Civilized City designation represent rare and valuable intangible assets that signal comprehensive urban quality-encompassing infrastructure excellence, service standards, and governance capacity; thereby reducing information asymmetries that typically constrain tourism decision-making (Carvalho et al., 2018; Horng & Tsai, 2012; Peters et al., 2011). This signaling effect creates a virtuous cycle: as cities pursue brand certification, they systematically enhance both tangible infrastructure and intangible service environments, generating spillover benefits that extend beyond the original branding objectives into tourism attractiveness and competitiveness (Cerqua, 2017; Yang et al., 2016). The certification process itself drives cities to adopt sustainable development practices (from environmental protection to resource conservation) that align organically with sustainable tourism principles, while the resulting brand recognition amplifies media exposure and consumer trust, creating powerful network effects that multiply tourism appeal. Through this integrated process, city branding transcends traditional marketing to become a comprehensive development strategy that simultaneously builds urban competitive advantages and tourism sustainability foundations. Based on the above theoretical analysis, this study proposes the following hypotheses:

H1: City branding promotes sustainable tourism development levels.

Understanding how city branding influences sustainable tourism development requires a deep examination of its underlying mechanisms. This complex process is not a simple linear relationship, but rather operates through the interactive effects of multiple theoretical mechanisms. Based on existing theoretical foundations and empirical observations, we identify two core transmission pathways that together constitute the theoretical framework for how city branding promotes sustainable tourism development.

The first pathway manifests through city branding's enhancement of tourism market attractiveness via signaling mechanisms. In the tourism decision-making process, potential tourists face significant information asymmetry problems. They find it difficult to accurately assess destination service quality, safety levels, and overall experiences before actual visits. According to the core tenets of signaling theory, high-quality service providers actively seek credible ways to convey quality information to the market, and certification by authoritative institutions represents precisely such an effective signaling mechanism (Auriol & Schilizzi, 2015; Boiral, 2012). The National Civilized City selection serves as a national-level authoritative certification system that not only represents rigorous evaluation standards but also carries government credibility endorsement (X. Zhao et al., 2025). When a city receives this designation, it transmits multiple positive signals to potential tourists: beautiful urban environments, good public order, comprehensive service systems, and high citizen quality. These signals effectively reduce tourists' perceived risks and decision costs, enhancing the city's attractiveness and competitive advantage in tourism destination selection. More importantly, this signaling effect possesses sustainability because the evaluation and re-assessment mechanisms of civilized cities ensure that cities must continuously maintain high standards, thereby providing stable quality assurance for sustainable tourism development.

The second pathway operates through resource integration mechanisms that stimulate tourism entrepreneurial vitality. Resource-based theory emphasizes that competitive advantages of enterprises or organizations stem from their unique resources, particularly those strategic resources that are valuable, rare, inimitable, and non-substitutable (Acedo et al., 2006; Kozlenkova et al., 2014). City branding, especially authoritative certifications like National Civilized City designation, precisely embodies these characteristics. It is not only a scarce intangible asset but also a strategic resource capable of generating multiple effects. First, city branding enhances cities' resource attraction capabilities, including capital, talent, technology, and other factor resources, providing richer factor support for tourism industry development. Second, city branding helps optimize the business environment, as governments often increase public service provision, simplify administrative approval processes, and improve infrastructure construction to maintain and enhance city brand image (X. Zhao et al., 2025). These measures directly reduce tourism enterprises' operational costs and market entry barriers. Finally, city branding creates positive market expectations, enhancing investors' and entrepreneurs' confidence in local tourism market prospects and stimulating tourism entrepreneurial enthusiasm. This enhancement of entrepreneurial vitality manifests not only in the growth of new tourism enterprises but also in the emergence of tourism product and service innovations, injecting sustained endogenous momentum into sustainable tourism development. Based on the above theoretical analysis, this study proposes the following hypotheses:

H2: City branding promotes sustainable tourism development by enhancing tourism market attractiveness.

H3: City branding promotes sustainable tourism development by stimulating tourism entrepreneurial vitality.

2 Methods

2.1 Data sources

This study selects panel data from 281 Chinese cities from 2002–2022, aiming to systematically evaluate the impact of National Civilized City selection on sustainable tourism development. Data sources primarily rely on authoritative statistical materials. The selection times and city lists for National Civilized Cities were compiled from official documents published by China's Government Civilization Committee, with designated National Civilized Cities set as the experimental group and remaining cities as the control group. Control variable data mainly comes from the annual China City Statistical Yearbook. To ensure comprehensiveness and accuracy, tourism enterprise-related data was simultaneously obtained from the Qichacha website, which provided detailed registration and operational information about tourism enterprises in various regions. Tourism resource endowment data was manually collected through relevant professional websites, comprehensively organizing and evaluating tourism resources in each city. Additionally, other control variable data was sourced from the CEIC database, regional statistical yearbooks, statistical bulletins, and the EPS data platform and other official statistical materials. These diverse data sources provided authoritative, systematic, and comprehensive data support for this research, effectively ensuring the reliability of empirical results.

2.2 Variable definitions

Regarding variable settings, this study's core explanatory variable is the National Civilized City selection dummy variable (NCC), which takes a value of 1 when a city receives or maintains National Civilized City status in a specific year, and 0 when it has not received the title or loses it due to failing re-evaluation. Considering research object comparability and data availability, districts and counties of municipalities, prefecture-level cities in Tibet Autonomous Region, and city samples that underwent major administrative division adjustments were excluded during the research process. Ultimately, 129 cities were identified as the treatment group to ensure the robustness and representativeness of research results. The explained variable is set as the sustainable tourism development level (TQD). Based on existing research findings and data availability (Juan Wang et al., 2025; Tang, 2022), this study employs the entropy weight-TOPSIS method to construct a comprehensive evaluation system. This method effectively considers the objective weights of various indicators, comprehensively reflecting the multidimensional characteristics of tourism development, avoiding the potential one-sidedness of single-indicator evaluations, and enhancing the persuasiveness and scientific nature of research conclusions.

The conceptualization of sustainable tourism development level draws upon the United Nations Sustainable Development Goals (SDGs) framework, emphasizing the coordinated integration of economic growth, social progress, environmental protection, and cultural preservation. We define sustainable tourism development level as a city's comprehensive capacity to achieve balanced economic, social, environmental, and cultural benefits throughout its tourism development process. To operationalize this multifaceted concept, we construct a five-dimensional evaluation system grounded in established theoretical foundations. The innovation dimension, rooted in endogenous growth theory, captures tourism industry technological advancement and human capital through indicators including tourism-related patent applications and industry employment, reflecting the sector's innovation momentum and talent base. The coordination dimension, informed by industrial linkage theory, measures tourism-regional economic integration through tourism revenue's share of tertiary industry output and tertiary sector's GDP contribution, demonstrating inter-sectoral synergies. The green dimension, based on Environmental Kuznets Curve theory, assesses environmental sustainability through forest coverage rates, waste treatment efficiency, and urban green space availability, capturing the ecological foundation for tourism development. The openness dimension, grounded in comparative advantage theory, evaluates internationalization levels via inbound tourist volumes and foreign exchange earnings, reflecting global competitiveness. Finally, the sharing dimension, informed by inclusive growth theory, measures development outcome accessibility through tourism investment, travel agencies, star-rated hotels, and A-level attractions, emphasizing equitable benefit distribution. This framework employs the entropy weight method for objective indicator weighting, ensuring scientific rigor while avoiding subjective bias in evaluation outcomes. The specific indicators and their detailed specifications are presented in Tab. 1.

To deeply analyze the internal mechanisms through which National Civilized City selection influences sustainable tourism development, this study establishes two mechanism variables: tourism market attractiveness and tourism entrepreneurial vitality. Tourism market attractiveness follows existing research approaches (Wu et al., 2023), measured using the tourism arrival ratio (TAT), an indicator that intuitively and accurately reflects a region's ability to attract tourists and the market conversion efficiency of tourism resources. Tourism entrepreneurial vitality selects the number

Tab. 1: Sustainable tourism development level evaluation indicator system

Target layer	Primary indicator	Secondary indicator	Unit
Sustainable tourism development level	Innovation	Tourism-related patent applications	Number
		Tourism industry employees	People
		Number of tourism colleges	Institutions
		Tourism college students	People
	Coordination	Tourism total revenue as proportion of tertiary industry output	%
		Tertiary industry added value as proportion of GDP	%
	Green	Forest coverage rate	%
		Urban domestic waste treatment rate	%
		Per capita park green space area	Hectares
	Openness	Inbound tourist numbers	10,000 person-times
		International tourism foreign exchange earnings	10,000 USD
	Sharing	Tourism fixed asset investment	10,000 yuan
		Number of travel agencies	Number
		Number of star-rated hotels	Number
		Number of A-level and above tourist attractions	Number

Source: own

of newly established tourism enterprises (TEA) as a measurement indicator, comprehensively reflecting regional tourism market vitality and innovation momentum, and demonstrating the promotional role of city branding in cultivating tourism market entities. Simultaneously, to comprehensively control interference from other factors and ensure research result accuracy, this study also incorporates multiple control variables: urban economic development level (PGDP), measured using per capita regional GDP; population size (POP), represented by the natural logarithm of urban resident population; infrastructure level (INFRA), measured by the ratio of urban road area to built-up area; industrial structure advancement (IND), measured by tertiary industry added value as a proportion of GDP; environmental regulation intensity (ENV), represented by industrial pollution control investment as a proportion of industrial added value; government intervention degree (GOV), measured by general public budget expenditure as a proportion of GDP; openness level (OPEN), measured by actual utilized foreign direct investment as a proportion of GDP; and technological innovation capability (TECH), measured by the natural logarithm of patent grants.

2.3 Model specification

Regarding model specification, this study constructs a multiple-period difference-in-differences model to evaluate the impact effect of National Civilized City selection on sustainable tourism development. This method's advantage lies in effectively controlling unobservable time fixed effects and individual fixed effects while overcoming the limitations of traditional difference-in-differences models when handling situations where policy implementation timing differs, particularly suitable for the real-world situation in this study where different cities received National Civilized City designation at different times. The basic econometric model is specified as:

$$TQD_{it} = \beta_0 + \beta_1 NCC_{it} + \sum \rho X_{it} + \delta_t + + \mu_i + \epsilon_{it} \quad (1)$$

where: TQD_{it} – the sustainable tourism development level of city i in year t; NCC_{it} – the core explanatory variable taking value 1 when a city receives National Civilized City designation and 0 otherwise; X_{it} – a vector of control variables; δ_t and μ_i – time fixed effects and city fixed effects, respectively, controlling for city characteristics that do not change over time and time trends affecting all cities; ε_{it} – the random disturbance term; coefficient β_1 – measures the net effect of National Civilized City selection on sustainable tourism development. If β_1 is significantly positive, it indicates that city branding can enhance sustainable tourism development levels, providing important empirical evidence for positive interactions between urban civilization construction and sustainable tourism development.

3 Results

3.1 Description of sustainable tourism development levels

Descriptive statistical analysis reveals significant regional heterogeneity in sustainable tourism development levels across China (Fig. 1). Eastern regions demonstrate clear leadership, with Guangdong Province achieving the highest average level (0.4969), followed by Jiangsu (0.4820) and Shandong (0.4826), reflecting their advanced economic development and tourism infrastructure. Central regions exhibit moderate performance, led by Anhui Province (0.4650), benefiting from policy initiatives at renowned scenic areas like Huangshan. Western regions show notable variation, with Chongqing outperforming most eastern provinces (0.4865) due to its municipal status and unique geographic advantages, while Sichuan (0.4707) and Guangxi (0.4651) leverage their rich natural and cultural resources. The northeastern provinces cluster around similar levels, with Liaoning slightly ahead (0.4622). Overall, China's sustainable tourism development exhibits a spatial pattern of "eastern leadership with selective western breakthroughs", closely linked to regional economic development, transportation accessibility, and tourism resource endowments.

3.2 Baseline model results

To investigate the impact of National Civilized City selection on sustainable tourism development levels, we first conducted

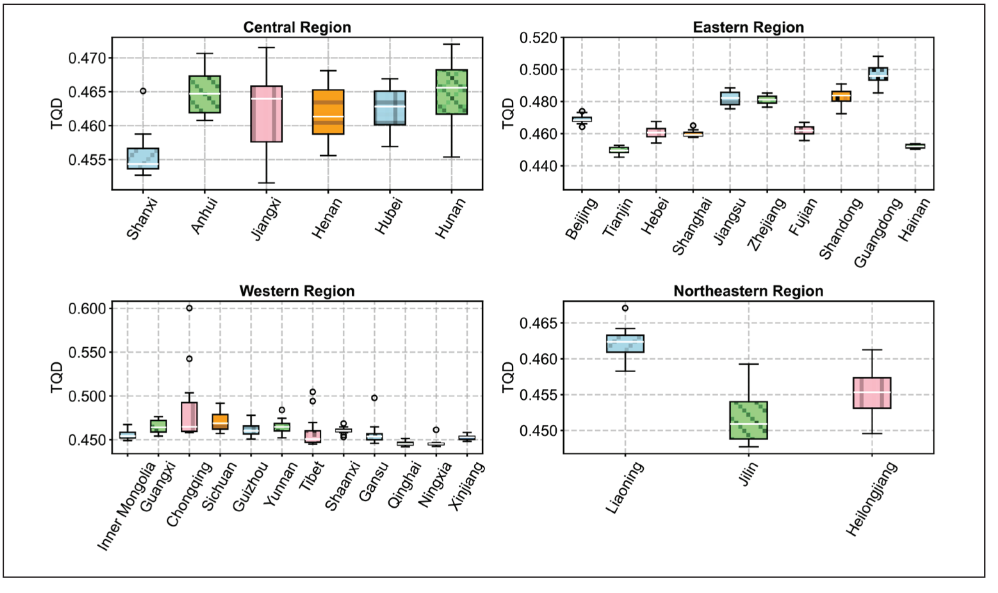

Fig. 1: Status of sustainable tourism development levels

Source: own

Tab. 2: Baseline regression results

Variable	(1)	(2)
	TQD	TQD
NCC	0.0326***	0.0284**
	(2.9100)	(2.5300)
PGDP		0.0147**
		(2.4200)
POP		−0.0008
		(−0.7500)
INFRA		0.0235***
		(3.0500)
IND		0.0128**
		(2.4700)
ENV		0.0094*
		(1.9500)
GOV		0.0076
		(1.2900)
OPEN		0.0182**
		(2.1600)
TECH		0.0156**
		(2.1000)
City	Yes	Yes
Year	Yes	Yes
N	5,620	5,620
R^2	0.5370	0.6470

Note: ***, **, and * indicate significance at the 1, 5, and 10% levels, respectively; values in parentheses are *t*-statistics.

Source: own

baseline regression analysis, with results shown in Tab. 2. Column 1 controls only for city and year fixed effects, while column 2 adds control variables on this basis. As shown in Tab. 2, the coefficient of National Civilized City (NCC) is positive and statistically significant under different specifications, indicating that receiving National Civilized City designation has a positive impact on enhancing urban sustainable tourism development levels. Specifically, when controlling only for fixed effects (column 1), receiving National Civilized City designation increases sustainable tourism development levels by an average of 0.0326 units; after adding other control variables (column 2), although the coefficient decreases somewhat (0.0284), it remains significant at the 5% level, providing strong empirical support for *H1*. This result reflects that the positive impact of National Civilized City selection remains robust after controlling for potential confounding factors, demonstrating that cities can enhance their attractiveness to tourists and tourism investment through authoritative branding certification.

3.3 Robustness tests

To verify the reliability of our research conclusions, we first tested for parallel trends. In 2005, the first batch of "National Civilized Cities" were selected and commended. We set 2004 as the base period and unified the fifth period after policy implementation and subsequent periods as the fifth period, constructing the following dynamic equation:

$$TQD_{it} = \alpha + \sum_{k=-4}^{5} \beta_k NCC_{i,t+k} + \gamma X_{it} + + \delta_t + \mu_i + \epsilon_{it} \quad (2)$$

As shown in Fig. 2, before policy implementation, the treatment and control groups exhibited essentially consistent trends in sustainable tourism development levels, with no significant differences between groups. This phenomenon indicates that before receiving National Civilized City designation, treatment and control groups followed similar development trajectories, meeting the basic requirements of the parallel trends assumption and effectively eliminating the possibility of reverse causality. After policy implementation, the treatment group's sustainable tourism development levels began to significantly exceed the control group's, with the gap gradually widening over time, providing intuitive evidence that National Civilized City selection promotes sustainable tourism development.

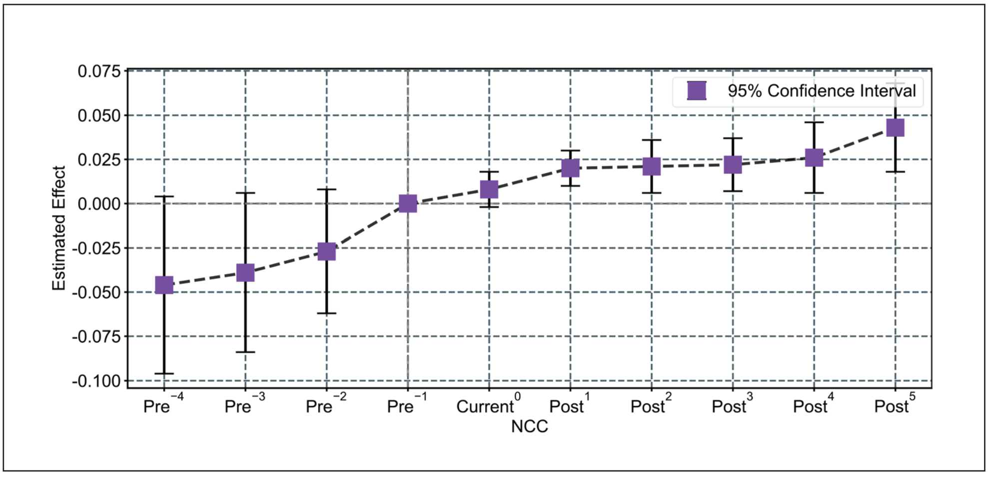

Fig. 2: **Parallel trends test**

Source: own

Next, we conducted a placebo test, advancing the National Civilized City policy implementation time by two years and re-estimating the policy effect. As shown in Fig. 3, before and after the fictional policy time point, changes in differences between treatment and control groups' sustainable tourism development levels were not significant, with the coefficient of the pseudo-treatment effect statistically insignificant, further verifying that the policy effect found in the baseline model was not driven by other factors or trends.

Third, we replaced the measurement method for the explained variable, using only the entropy weight method rather than the entropy weight-TOPSIS method to measure sustainable tourism development levels, testing result sensitivity to indicator construction methods. Tab. 3 shows regression results after replacing the proxy variable, with the coefficient of National Civilized City selection remaining positive and significant (0.0275), indicating that our research conclusions have strong robustness.

Fourth, we adopted propensity score matching combined with difference-in-differences (PSM-DID) to further mitigate selection bias issues. Tab. 4 reports PSM-DID estimation results, with the coefficient of National Civilized

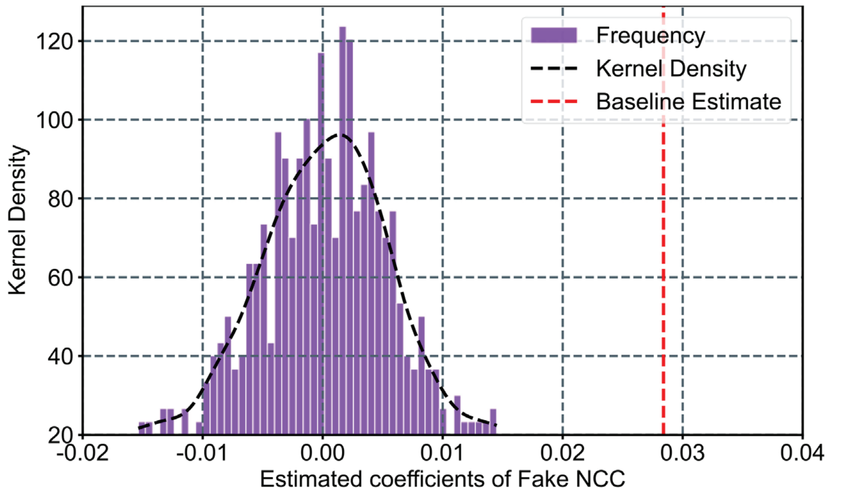

Fig. 3: Placebo test

Source: own

Tab. 3: Regression results with alternative proxy variables

Variable	(1)	(2)
	TQD	TQD
NCC	0.0318***	0.0275**
	(2.7200)	(2.3500)
Controls	No	Yes
City	Yes	Yes
Year	Yes	Yes
N	5,620	5,620
R^2	0.5290	0.6350

Note: ***, **, and * indicate significance at the 1, 5, and 10% levels, respectively; values in parentheses are *t*-statistics.

Source: own

City selection remaining positive and significant under both nearest neighbor matching (0.0258) and kernel matching (0.0262) methods, though slightly smaller than the baseline model, indicating more accurate policy effect estimation after considering systematic differences in observable characteristics. Finally, to exclude interference from other concurrent policies, we added dummy variables for concurrent policies of new urbanization demonstration zones and smart city pilots to the model. As shown in column 3 of Tab. 4, after controlling for these concurrent policies, the coefficient of National Civilized City selection (0.0271) remained

Tab. 4: PSM-DID test and exclusion of concurrent policy interference

Variable	(1)	(2)	(3)
	Nearest neighbor matching	Kernel matching	Controlling for concurrent policies
	TQD	TQD	TQD
NCC	0.0258**	0.0262**	0.0271**
	(2.2300)	(2.4900)	(2.4100)
Controls	Yes	Yes	Yes
City	Yes	Yes	Yes
Year	Yes	Yes	Yes
N	4,328	4,876	5,620
R^2	0.6350	0.6420	0.6520

Note: ***, **, and * indicate significance at the 1, 5, and 10% levels, respectively; values in parentheses are *t*-statistics.

Source: own

significantly positive with little change in value, indicating that our identified policy effect indeed comes from National Civilized City selection rather than other urban development policies. Our research conclusions are strongly supported through this series of robustness tests.

3.4 Heterogeneity analysis

To explore in depth the differentiated impact of National Civilized City selection on sustainable tourism development in different types of cities, this paper conducted heterogeneity analysis from the dimensions of industrial structure and digital infrastructure, with results shown in Tab. 5. In the industrial structure dimension, using tertiary industry added value as a proportion of GDP above the median as the standard, the sample was divided into service-led cities and non-service-led cities. Results from columns 1–2 show that the enhancement effect

Tab. 5: Heterogeneity test results

Variable	(1)	(2)	(3)	(4)
	Service-led	Non-service-led	High digitalization	Low digitalization
	TQD	TQD	TQD	TQD
NCC	0.0327***	0.0235**	0.0342***	0.0218**
	(3.1400)	(2.1000)	(3.4900)	(2.0400)
Controls	Yes	Yes	Yes	Yes
City	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
R^2	0.6580	0.6290	0.6640	0.6310
Empirical *p*-value	0.0420		0.0370	

Note: ***, **, and * indicate significance at the 1, 5, and 10% levels, respectively; values in parentheses are *t*-statistics.

Source: own

of National Civilized City selection on sustainable tourism development levels is significantly higher in service-led cities (0.0327) than in non-service-led cities (0.0235), with the *p*-value for the difference between the two groups' coefficients at 0.042, significant at the 5% level. This result indicates that cities with better service industry development foundations can more fully utilize the National Civilized City brand effect, achieving coordinated development between tourism and related service industries, forming a virtuous interactive cycle.

In the digital infrastructure dimension, using whether the number of internet broadband access users per 10,000 people is above the median as the standard, the sample was divided into high-digitalization cities and low-digitalization cities. Results from columns 3–4 show that the promotional effect of National Civilized City selection on sustainable tourism development levels is significantly stronger in high-digitalization cities (0.0342) than in low-digitalization cities (0.0218), with the *p*-value for the difference between the two groups' coefficients at 0.037, significant at the 5% level. This indicates that cities with well-developed digital infrastructure can better leverage information technology to amplify the civilized city brand effect, such as through promoting city image via digital platforms, providing smart tourism services, and optimizing tourism management, thereby more effectively enhancing sustainable tourism development levels. These heterogeneity analysis results provide important references for cities to formulate tourism development strategies suited to local conditions, with different types of cities needing to fully leverage National Civilized City brand value based on their own conditions to promote sustainable tourism development.

3.5 Mechanism tests

Based on the theoretical analysis above, this study further explores the specific mechanisms through which National Civilized City selection influences sustainable tourism development, focusing on examining two potential channels: tourism market attractiveness and tourism entrepreneurial vitality, with related results shown in Tab. 6. Columns 1–2 show the impact of National Civilized City selection (NCC) on tourism market attractiveness (TAT). Results show that after controlling for relevant factors, receiving National Civilized City designation increases a city's tourism arrival ratio by an average of 0.1642%, significant at the 1% level. This indicates that National Civilized City selection significantly enhances cities' attractiveness to tourists and expands tourism market scale by improving city recognition, enhancing city

Tab. 6: Mechanism analysis results

Variable	(1)	(2)	(3)	(4)
	TAT	TQD	TEA	TQD
NCC	0.1642***	0.0264**	0.1847**	0.0251**
	(3.6300)	(2.3500)	(2.3600)	(2.2400)
TAT		0.0124***		
		(3.3500)		
TEA				0.0181***
				(2.8300)
Controls	Yes	Yes	Yes	Yes
City	Yes	Yes	Yes	Yes
Year	Yes	Yes	Yes	Yes
N	5,620	5,620	5,620	5,620
R^2	0.6320	0.6390	0.6240	0.6350

Note: ***, **, and * indicate significance at the 1, 5, and 10% levels, respectively; values in parentheses are *t*-statistics.

Source: own

image, and optimizing urban environment. Columns 3–4 show the impact of National Civilized City selection on tourism entrepreneurial vitality (TEA), with results indicating that receiving National Civilized City designation increases the number of newly established tourism enterprises by an average of 18.47, significant at the 5% level. This confirms that National Civilized City selection effectively promotes the cultivation of tourism market entities and activates tourism entrepreneurship by improving the urban business environment, boosting market confidence, and enhancing tourism prospect expectations.

Furthermore, we analyzed the impact of tourism market attractiveness and tourism entrepreneurial vitality on sustainable tourism development levels. Results show that for each 1% increase in tourism arrival ratio (TAT), sustainable tourism development levels increase by an average of 0.0124 units; for each 1% increase in newly established tourism enterprises (TEA), sustainable tourism development levels increase by an average of 0.0181 units, with coefficients for both variables statistically significant. Combining these results, we can confirm that National Civilized City selection significantly promotes sustainable tourism development level enhancement by improving tourism market attractiveness and promoting tourism entrepreneurial vitality. This mechanism analysis not only provides empirical evidence for understanding the inherent pathways of city branding in promoting tourism development but also offers important references for related policy formulation and practice.

Synthesizing the mechanism analysis results, we can confirm that National Civilized City selection promotes sustainable tourism development level enhancement through two pathways: "improving city image → enhancing tourism attractiveness → expanding tourism market scale" and "optimizing business environment → stimulating tourism entrepreneurial vitality → strengthening industry development momentum", forming a "dual-wheel drive" mechanism. These findings provide robust empirical validation for *H2* and *H3*, demonstrating that city branding operates through both demand-side market attraction and supply-side entrepreneurial stimulation mechanisms. This dual-pathway mechanism reveals the intrinsic connection between urban soft power construction and industrial development, indicating that civilized city creation should be incorporated into comprehensive urban industrial development strategies rather than treated as standalone image projects.

4 Discussion

The positive relationship between city branding and sustainable tourism development observed in our study engages with and extends existing international scholarship on place branding effectiveness while addressing critical gaps in mechanistic understanding. Our empirical findings regarding the significant impact of National Civilized City designation on tourism development validate the broader theoretical proposition that place branding substantially influences visitors' responses, leading to improved destination image, trust, value, and loyalty (Ozer et al., 2025). However, our contribution transcends mere validation by empirically identifying and quantifying specific transmission mechanisms, tourism market attractiveness and entrepreneurial vitality, that previous research has largely treated as theoretical constructs rather than measurable pathways. This mechanistic precision addresses a critical gap in the literature where scholars have acknowledged the importance of place branding for tourism outcomes without adequately explaining how these effects materialize in practice through identifiable causal chains (Rinaldi et al., 2021). Our comprehensive evaluation framework for sustainable tourism development, encompassing innovation, coordination, green development, openness, and sharing dimensions, aligns with contemporary perspectives that recognize the necessity of integrating economic vitality, social equity, and environmental responsibility in destination branding strategies (Khater et al., 2024). The multidimensional nature of our measurement system resonates with research identifying key clustering attributes for sustainable destinations, including cultural and natural attributes, seasonality considerations, and tourist profiles, while extending beyond descriptive categorization to provide quantitative assessment capabilities that enable more rigorous empirical testing (Ivars-Baidal et al., 2023; Schianetz & Kavanagh, 2008). This advancement in measurement methodology represents an important step toward operationalizing abstract sustainability concepts into tangible evaluation frameworks that can guide destination management and policy formulation.

The heterogeneous effects we identified across different urban typologies engage with emerging theoretical frameworks that emphasize the contextual nature of place branding effectiveness and challenge conventional assumptions about universal applicability of branding strategies. Our finding that service-led and highly digitalized cities experience substantially stronger benefits from civilized city designation supports scholarship highlighting the role of boundary spanners and local capacity in mediating between global branding strategies and local implementation contexts (Rinaldi et al., 2021). This heterogeneity challenges universalist assumptions about place branding effectiveness while supporting research demonstrating that stakeholders from local state, destinations, businesses, and communities negotiate influence and legitimacy in the place branding process, with institutional arrangements and cultural capital playing significant mediating roles (Leiras et al., 2025; Reynolds et al., 2022). Our results extend this understanding by quantifying how existing urban characteristics, particularly service sector development and digital infrastructure, moderate branding effectiveness, suggesting that successful place branding requires not merely strategic communication but foundational institutional and economic capabilities that enable cities to leverage brand value effectively (Hossain et al., 2025). The implications of this finding extend beyond China's institutional context to inform broader debates about the conditions under which place branding initiatives can generate meaningful sustainability outcomes rather than merely appropriating sustainability discourse for competitive advantage. The cultural heritage aspects emphasized in our analysis of National Civilized City designation complement international research on heritage-based place branding, which explores how cultural assets can be leveraged for sustainable tourism development through network formation and collaborative governance mechanisms (Pai et al., 2025). However, our findings advance this conversation by demonstrating that culturally-oriented city brands can effectively support sustainable tourism development across different institutional contexts, provided they are backed by genuine improvements in urban governance and service quality rather than superficial image management exercises that prioritize symbolic politics over substantive policy interventions (Bruin et al., 2025; Gonzalez & Gale, 2023). This distinction between authentic and performative branding represents an important contribution to ongoing debates about the relationship between city branding and sustainable development outcomes.

Conclusions

Main findings. Based on panel data from 281 Chinese cities from 2002–2022, this study systematically examines the impact of National Civilized City selection on sustainable tourism development and its mechanisms using a multiple-period difference-in-differences model. Our empirical analysis reveals three key findings that challenge conventional wisdom about city branding effectiveness. First, National Civilized City designation significantly enhances urban sustainable tourism development levels by an average of 0.0284 units, a finding that remains robust across multiple specification tests and alternative methodologies. This effect represents not merely statistical significance but practical importance, demonstrating that authoritative city certifications can indeed translate abstract brand value into measurable tourism outcomes. Second, we identify and empirically validate two distinct transmission mechanisms through which city branding operates: enhancing tourism market attractiveness and stimulating tourism entrepreneurial vitality. This dual-pathway mechanism forms what we term a "virtuous amplification cycle", where improved city image attracts more tourists while simultaneously creating business opportunities that further enhance the city's tourism appeal. Third, the promotional effects demonstrate significant heterogeneity across urban typologies, with service-dominated and highly digitalized cities experiencing substantially stronger benefits from civilized city designation. This heterogeneity suggests that city branding effectiveness is not universal but contingent upon existing urban capabilities and institutional arrangements.

Theoretical contributions. This study advances theoretical understanding across multiple domains, offering contributions that extend beyond traditional disciplinary boundaries. From a resource-based theory perspective, we demonstrate that city brands function as unique intangible resources capable of generating sustainable competitive advantages through the VRIO framework (valuable, rare, inimitable, and organizationally embedded), they are valuable in attracting tourists and investment, rare due to stringent selection criteria, inimitable because of their institutional backing, and organizationally embedded within urban governance structures. Our signaling theory contributions lie in empirically validating how authoritative city certifications serve as quality signals that reduce information asymmetry between destinations and potential tourists, extending signaling theory beyond traditional corporate applications to public sector branding contexts. From brand spillover effects theory, we reveal how city brands transform into tourism brand assets through extension mechanisms, expanding brand theory's applicable boundaries by showing how general city reputation can be channeled into sector-specific advantages. Additionally, our construction of a multidimensional sustainable tourism development evaluation system provides new methodological contributions for quantitatively measuring tourism sustainability, integrating innovation, coordination, green development, openness, and sharing dimensions to offer a more nuanced assessment tool than existing single-indicator approaches.

Practical implications. The practical implications of this research extend far beyond academic interest, offering strategic insights that can reshape how urban managers approach city branding and tourism development. Our findings demonstrate that city branding represents more than symbolic politics or image management, it constitutes a powerful economic development tool capable of generating measurable returns on investment. Urban managers should therefore integrate branding activities such as civilized city creation into comprehensive urban industrial development strategies rather than treating them as standalone public relations exercises. The identification of heterogeneous effects across city types suggests that successful branding strategies must be tailored to local contexts and capabilities. Service-dominated cities should focus on leveraging synergies between tourism and other service industries, exploring opportunities for industrial chain extension and value chain enhancement that maximize cross-sectoral spillovers. Highly digitalized cities should capitalize on their technological advantages to accelerate smart tourism development, using digital platforms to amplify branding effects and enhance visitor experiences through innovation. For cities lacking these comparative advantages, our findings suggest that building foundational capabilities in service sector development or digital infrastructure may be prerequisites for maximizing branding investment returns. Additionally, the dual-mechanism nature of branding effects, operating through both market attractiveness and entrepreneurial vitality, implies that successful city branding requires coordinated attention to both demand-side marketing and supply-side business environment optimization.

Research limitations and future directions. While this study provides robust evidence for city branding effects on sustainable tourism development, several limitations suggest important directions for future research. First, our analysis focuses exclusively on Chinese cities, limiting the international generalizability of our findings. The institutional context of China's National Civilized City program, with its strong government backing and comprehensive evaluation criteria, may not translate directly to other national contexts where city branding operates through different mechanisms or institutional arrangements. Future research should examine similar phenomena across diverse institutional and cultural contexts to establish the universality of our identified mechanisms.

Second, our measurement of sustainable tourism development, while comprehensive in its five-dimensional framework, relies primarily on quantitative indicators that may not fully capture qualitative aspects of sustainability. Although our evaluation system integrates innovation, coordination, green development, openness, and sharing dimensions, certain critical elements, such as local community satisfaction, cultural authenticity preservation, and long-term environmental resilience remain challenging to quantify using available statistical data. Future research could complement our quantitative approach with qualitative assessments that incorporate stakeholder perspectives, particularly from local residents and small-scale tourism operators whose voices are often marginalized in large-scale statistical analyses. Mixed-methods approaches combining panel data analysis with in-depth case studies could provide richer understanding of how city branding initiatives actually affect different community groups and whether the observed tourism development genuinely aligns with sustainability principles beyond measurable indicators.

Third, our study examines a relatively short time horizon following National Civilized City designation, which may not adequately capture long-term dynamics and potential diminishing returns of branding effects. While our dataset spans two decades, the temporal lag between certification and observable impacts, combined with periodic re-evaluations of civilized city status, creates complexity in distinguishing between initial branding effects and sustained impacts. The possibility exists that branding effects may decay over time as the novelty wears off or as competing cities acquire similar designations, potentially diluting the competitive advantage initially gained. Longitudinal research extending beyond our observation period could examine whether the positive effects we identified persist, strengthen, or weaken over extended timeframes, and whether cities must continuously reinvest in brand maintenance to sustain tourism development benefits. Additionally, future studies could investigate potential non-linear relationships between branding duration and tourism outcomes, exploring whether optimal branding lifecycles exist.

Fourth, our quantitative methodology, while rigorous for establishing causal relationships, provides limited insight into the micro-level processes through which branding effects manifest in practice. Future research combining quantitative analysis with qualitative case studies could illuminate the behavioral mechanisms underlying our statistical findings, particularly regarding how tourists, entrepreneurs, and other stakeholders actually respond to city branding initiatives.

References

Acedo, F. J., Barroso, C., & Galan, J. L. (2006). The resource-based theory: Dissemination and main trends. *Strategic Management Journal*, *27*(7), 621–636. https://doi.org/10.1002/smj.532

Andersson, I., & James, L. (2018). Altruism or entrepreneurialism? The co-evolution of green place branding and policy tourism in Växjö, Sweden. *Urban Studies*, *55*(15), 3437–3453. https://doi.org/10.1177/0042098017749471

Artal-Tur, A., Briones-Peñalver, A. J., Bernal-Conesa, J. A., & Martínez-Salgado, O. (2019). Rural community tourism and sustainable advantages in Nicaragua. *International Journal of Contemporary Hospitality Management*, *31*(6), 2232–2252. https://doi.org/10.1108/IJCHM-05-2018-0429

Auriol, E., & Schilizzi, S. G. M. (2015). Quality signaling through certification in developing countries. *Journal of Development Economics*, *116*, 105–121. https://doi.org/10.1016/j.jdeveco.2015.03.007

Boiral, O. (2012). ISO certificates as organizational degrees? Beyond the rational myths of the certification process. *Organization Studies*, *33*(5–6), 633–654. https://doi.org/10.1177/0170840612443622

Bruin, R., Cooper, E., Lamers, M., & Blair, B. (2025). Sustaining the brand and branding sustainability: Landscape as homeland in Greenland's visual marketing. *Journal of Sustainable Tourism*, *33*(11), 2484–2503. https://doi.org/10.1080/09669582.2024.2400572

Carvalho, J. M. S., Costa, R. V., Marnoto, S., Sousa, C. A. A., & Vieira, J. C. (2018). Toward a resource-based view of city quality: A new framework. *Growth and Change*, *49*(2), 266–285. https://doi.org/10.1111/grow.12237

Cerqua, A. (2017). The signalling effect of eco-labels in modern coastal tourism. *Journal of Sustainable Tourism*, *25*(8), 1159–1180. https://doi.org/10.1080/09669582.2016.1257014

Chen, N. (Chris), Dwyer, L., & Firth, T. (2014). Effect of dimensions of place attachment on residents' word-of-mouth behavior. *Tourism Geographies*, *16*(5), 826–843. https://doi.org/10.1080/14616688.2014.915877

Eslami, S., Khalifah, Z., Mardani, A., Streimikiene, D., & Han, H. (2019). Community attachment, tourism impacts, quality of life and residents' support for sustainable tourism development. *Journal of Travel & Tourism Marketing*, *36*(9), 1061–1079. https://doi.org/10.1080/10548408.2019.1689224

Fok, K. W. K., & Law, W. W. Y. (2018). City re-imagined: Multi-stakeholder study on branding Hong Kong as a city of greenery. *Journal of Environmental Management*, *206*, 1039–1051. https://doi.org/10.1016/j.jenvman.2017.11.045

Gautam, V. (2023). Why local residents support sustainable tourism development? *Journal of Sustainable Tourism*, *31*(3), 877–893. https://doi.org/10.1080/09669582.2022.2082449

Gonzalez, R. L., & Gale, F. (2023). Sustainability as economic value pluralism: Implications for urban politics and policy. *Cities*, *134*, 104167. https://doi.org/10.1016/j.cities.2022.104167

Horng, J.-S., & Tsai, C.-T. (Simon). (2012). Constructing indicators of culinary tourism strategy: An application of resource-based theory. *Journal of Travel & Tourism Marketing*, *29*(8), 796–816. https://doi.org/10.1080/10548408.2012.730945

Hossain, M. I., Japutra, A., Islam, N., & Javed, A. (2025). Mindsets matter: Reimagining the destination brand experience-environmental behaviour link through a cognitive lens. *Journal of Environmental Management*, *394*, 127648. https://doi.org/10.1016/j.jenvman.2025.127648

Işık, C., Islam, H., Ongan, S., Şengöz, A., & Usanmaz, B. (2025). Artificial intelligence (AI) interacted ESG-based sustainable tourism: Economic insights. *Tourism Economics*, *32*(4). https://doi.org/10.1177/13548166251346306

Islam, Md. W., Ruhanen, L., & Ritchie, B. W. (2018). Tourism governance in protected areas: Investigating the application of the adaptive co-management approach. *Journal of Sustainable Tourism*, *26*(11), 1890–1908. https://doi.org/10.1080/09669582.2018.1526291

Ivars-Baidal, J. A., Vera-Rebollo, J. F., Perles-Ribes, J., Femenia-Serra, F., & Celdrán-Bernabeu, M. A. (2023). Sustainable tourism indicators: What's new within the smart city/destination approach? *Journal of Sustainable Tourism*, *31*(7), 1556–1582. https://doi.org/10.1080/09669582.2021.1876075

Johansson, M. (2012). Place branding and the imaginary: The politics of re-imagining a garden city. *Urban Studies*, *49*(16), 3611–3626. https://doi.org/10.1177/0042098012446991

Khater, M., Ibrahim, O., Sayed, M. N. E., & Faik, M. (2024). Legal frameworks for sustainable tourism: Balancing environmental conservation and economic development. *Current Issues in Tourism*, *28*(22), 3693–3714. https://doi.org/10.1080/13683500.2024.2404181

Kozlenkova, I. V., Samaha, S. A., & Palmatier, R. W. (2014). Resource-based theory in marketing. *Journal of the Academy of Marketing Science*, *42*(1), 1–21. https://doi.org/10.1007/s11747-013-0336-7

Leiras, A., Domínguez-Vila, T., & Magano, J. (2025). Accessible tourist destinations: A bifactorial image model of perceived image. *Tourism Management Perspectives*, *58*, 101400. https://doi.org/10.1016/j.tmp.2025.101400

Lucarelli, A. (2019). Constructing a typology of virtual city brand co-creation practices: An ecological approach. *Journal of Place Management and Development*, *12*(2), 227–247. https://doi.org/10.1108/JPMD-04-2018-0031

Ma, W., de Jong, M., Hoppe, T., & de Bruijne, M. (2021). From city promotion via city marketing to city branding: Examining urban strategies in 23 Chinese cities. *Cities*, *116*, 103269. https://doi.org/10.1016/j.cities.2021.103269

Magnusson, T., Karabag, S. F., Wigger, K., & Andersson, G. (2024). Sustainability transitions in tourism: On the transformation of a fragmented sector. *Tourism Geographies*, *26*(2), 157–172. https://doi.org/10.1080/14616688.2023.2291700

Mihalič, T., Šegota, T., Knežević Cvelbar, L., & Kuščer, K. (2016). The influence of the political environment and destination governance on sustainable tourism development: A study of Bled, Slovenia. *Journal of Sustainable Tourism*, *24*(11), 1489–1505. https://doi.org/10.1080/09669582.2015.1134557

Niedomysl, T., & Jonasson, M. (2012). Towards a theory of place marketing. *Journal of Place Management and Development*, *5*(3), 223–230. https://doi.org/10.1108/17538331211269639

Ozer, M., Can, A. S., Ekinci, Y., & Ozer, A. (2025). Leveraging cittaslow promotion on social media to enhance tourists' behavioral responses toward small urban tourism destinations. *Journal of Travel Research*, *65*(3), 812–832. https://doi.org/10.1177/00472875241311550

Pai, C.-H., Zhang, Y., Wang, Y.-L., Li, K., & Shang, Y. (2025). Current challenges and opportunities in cultural heritage preservation through sustainable tourism practices. *Current Issues in Tourism*, *29*(5), 1016–1034. https://doi.org/10.1080/13683500.2024.2443776

Peters, M., Siller, L., & Matzler, K. (2011). The resource-based and the market-based approaches to cultural tourism in alpine destinations. *Journal of Sustainable Tourism*, *19*(7), 877–893. https://doi.org/10.1080/09669582.2010.547198

Pham, K., Andereck, K. L., & Vogt, C. A. (2025). Stakeholders' involvement in an evidence-based sustainable tourism plan. *Journal of Sustainable Tourism*, *33*(4), 673–696. https://doi.org/10.1080/09669582.2023.2259117

Poudel, S., Nyaupane, G. P., & Budruk, M. (2016). Stakeholders' perspectives of sustainable tourism development: A new approach to measuring outcomes. *Journal of Travel Research*, *55*(4), 465–480. https://doi.org/10.1177/0047287514563166

Reynolds, L., Koenig-Lewis, N., Doering, H., & Peattie, K. (2022). Competing for legitimacy in the place branding process: (Re)negotiating the stakes. *Tourism Management*, *91*, 104532. https://doi.org/10.1016/j.tourman.2022.104532

Rinaldi, C., Giovanardi, M., & Lucarelli, A. (2021). Keeping a foot in both camps: Sustainability, city branding and boundary spanners. *Cities*, *115*, 103236. https://doi.org/10.1016/j.cities.2021.103236

Roxas, F. M. Y., Rivera, J. P. R., & Gutierrez, E. L. M. (2020). Mapping stakeholders' roles in governing sustainable tourism destinations. *Journal of Hospitality and Tourism Management*, *45*, 387–398. https://doi.org/10.1016/j.jhtm.2020.09.005

Schianetz, K., & Kavanagh, L. (2008). Sustainability indicators for tourism destinations: A complex adaptive systems approach using systemic indicator systems. *Journal of Sustainable Tourism*, *16*(6), 601–628. https://doi.org/10.1080/09669580802159651

Stylidis, D. (2022). Exploring resident-tourist interaction and its impact on tourists' destination image. *Journal of Travel Research*, *61*(1), 186–201. https://doi.org/10.1177/0047287520969861

Stylidis, D., Woosnam, K. M., & Tasci, A. D. A. (2022). The effect of resident-tourist interaction quality on destination image and loyalty. *Journal of Sustainable Tourism*, *30*(6), 1219–1239. https://doi.org/10.1080/09669582.2021.1918133

Tang, R. (2022). A study of the effects and mechanisms of the digital economy on high-quality tourism development: Evidence from the Yangtze River Delta in China. *Asia Pacific Journal of Tourism Research*, *27*(11), 1217–1232. https://doi.org/10.1080/10941665.2023.2174033

Wang, Juan, Wei, R., & Zhang, H. (2025). Marine policies' efficacy on tourism economy high-quality development: Empirical evidence from China's coastal region. *Journal of Quality Assurance in Hospitality & Tourism*, 1–31. https://doi.org/10.1080/1528008x.2025.2478592

Wang, Jing, Yin, X., Wu, Y., Du, S., & Du, Z. (2025). How can television dramas stimulate tourists' impulsive travel intentions? The mediating role of destination brand love. *Current Issues in Tourism*, 1–21. https://doi.org/10.1080/13683500.2025.2535562

Weaver, D. B., Moyle, B., & McLennan, C. J. (2022). The citizen within: Positioning local residents for sustainable tourism. *Journal of Sustainable Tourism*, *30*(4), 897–914. https://doi.org/10.1080/09669582.2021.1903017

Williams, H. A., Yuan, J. (Jessica), & Williams, R. L. (2019). Attributes of memorable gastro-tourists' experiences. *Journal of Hospitality & Tourism Research*, *43*(3), 327–348. https://doi.org/10.1177/1096348018804621

Wu, D., Ma, R., Hao, C., Feng, X., & Wu, Y. (2023). Spatial effect and impact mechanism of digital economy on tourism industry high-quality development in Chinese cities. *Economic Geography*, *43*(04), 229–240 (in Chinese). https://doi.org/10.15957/j.cnki.jjdl.2023.04.023

Wu, L., Taheri, B., Okumus, F., & Wang, S. (2024). The effects of host sincerity on tourists' perceived destination image. *The Service Industries Journal*, *44*(1–2), 83–104. https://doi.org/10.1080/02642069.2021.2011861

Yang, Y., Mueller, N. J., & Croes, R. R. (2016). Market accessibility and hotel prices in the Caribbean: The moderating effect of quality-signaling factors. *Tourism Management*, *56*, 40–51. https://doi.org/10.1016/j.tourman.2016.03.021

Zhao, L., Li, Q., & Sheng, Y. C. (2025). Environmental policies as engines for tourism eco-efficiency: Quasi-experimental evidence. *Journal of Sustainable Tourism*, *33*(9), 1890–1910. https://doi.org/10.1080/09669582.2024.2407491

Zhao, X., Lei, X., Yuan, H., & Zhang, C. (2025). Can urban "Gold Signboards" yield carbon reduction dividends? A quasi-natural experiment based on the "National Civilized City" selection. *Economics*, *19*(1). https://doi.org/10.1515/econ-2025-0150

All articles in the journal have been double-blind peer reviewed by a minimum of two independent experts in the relevant field. Authors are responsible for the linguistic accuracy of their manuscripts.

Journal title: **E+M EKONOMIE A MANAGEMENT**
E&M ECONOMICS AND MANAGEMENT

Address of the editorial office
Technical University of Liberec
Faculty of Economics
Editorial Office of the E&M Economics and Management journal
Voroněžská 13
461 17 Liberec 1
Czech Republic
Tel.: +420 485 352 345
E-mail: journal@tul.cz

The journal is published quarterly.

Subscription:
From 2024, the print format of the journal is suspended. The individual issues will be available for free in electronic form on the journal's website. The last print issue is issue 4, volume 27.

It is possible to order older issues only until present supplies are exhausted (20 EUR per issue).

Published by: Technical University of Liberec
Studentská 2, 461 17 Liberec 1, Czech Republic, ID no. 46747885

Production: Wolters Kluwer ČR, a. s.
U nákladového nádraží 10, 130 00 Praha 3, Czech Republic, ID no. 63077639

ISSN 1212-3609, ISSN (online) 2336-5064
www.ekonomie-management.cz