

Centre for Forensic
Accounting

UNIVERSITY OF BELGRADE
Faculty of Economics
and Business



Serbian Association
of Accountants and Auditors

Third Conference on Forensic Accounting

FROM ACCOUNTING MANIPULATION TO
SUSTAINABILITY AND ARTIFICIAL INTELLIGENCE:
NEW HORIZONS IN FORENSIC ACCOUNTING

Conference Proceedings

Editors

Dejan Malinić

Savka Vučković Milutinović

Belgrade, October 2026.

Third Conference on Forensic Accounting
**FROM ACCOUNTING MANIPULATION TO SUSTAINABILITY AND
ARTIFICIAL INTELLIGENCE: NEW HORIZONS IN FORENSIC ACCOUNTING**
Conference Proceedings

Publisher

University of Belgrade
Faculty of Economics and Business
Publishing Centre
Kamenička 6, Beograd
tel. +381 69 8066 416, +381 11 2633 884
e-mail: cid@ekof.bg.ac.rs
<http://cid.ekof.bg.ac.rs>

Co-publisher

Serbian Association of Accountants and Auditors
Njegoševa 19, Beograd
tel. +381 11 3239 444
www.srrs.rs

Dean of the Faculty of Economics and Business
Prof. dr Žaklina Stojanović

Director of the Publishing Centre
Prof. dr Đorđe Mitrović

For co-publisher
Dr Pero Škobić

Editors
Prof. dr. Dejan Malinić
Dr Savka Vučković Milutinović, Associate Professor

Cover design
MaxNova Creative – Beograd
www.maxnova.rs

Text editing and printing
JP „Službeni glasnik” – Beograd
www.slglasnik.com

Print run
150

Year
2026.

ISBN 978-86-403-1935-5

© 2026. University of Belgrade – Faculty of Economics and Business

All rights reserved. No part of this book may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage retrieval system, without permission from UNIVERSITY OF BELGRADE – FACULTY OF ECONOMICS AND BUSINESS.

TABLE OF CONTENTS

FOREWORD	VII
Douglas Cumming FRAUD, MANIPULATION, AND REGULATORY DESIGN: WHAT EVIDENCE FROM MATURE MARKETS CAN AND CANNOT TELL SERBIA AND THE REGION	1
Saeid Homayoun, Zabihollah Rezaee, Zoltán Oszkár SZÁNTÓ DO ESG-CENTRIC FIRMS ENGAGE IN FRAUDULENT ACTIVITIES?	39
Dejan Malinić, Nataša Džudović CONCEPTUAL FORENSIC FRAMEWORK FOR IDENTIFYING TRANSFER PRICING MANIPULATION RISK	69
Ivan Dionisijev, Zorica Bozhinovska Lazarevska, Monika Mitrevska PROFESSIONAL SKEPTICISM AND THE IDENTIFICATION OF IRREGULARITIES IN PUBLIC SECTOR AUDITS: EVIDENCE FROM STATE AUDITORS	97
Savka Vučković Milutinović, Tijana Ostojić INTERNAL AUDIT AS A VALUABLE ACTOR IN ADDRESSING FRAUD: WHERE DOES IT STAND IN SERBIA?	119
Miloš Božović MARKET MANIPULATION, MOMENTUM CRASHES, AND THE LIMITS OF PRICE-VOLUME SCREENING	141
Diana Dumitrescu, Nicolae Bobițan, Adriana Florina Popa ARTIFICIAL INTELLIGENCE, FORENSIC ACCOUNTING, AND TAX-CRIME INVESTIGATIONS: CROSS-COUNTRY EVIDENCE	155
Sofija Sekulić, Ana Lalević Filipović INSTITUTIONAL ENVIRONMENT AS A PREREQUISITE FOR THE DEVELOPMENT OF FORENSIC ACCOUNTING: A CLUSTER AND COMPARATIVE ANALYSIS OF EUROPEAN UNION AND WESTERN BALKAN COUNTRIES	171
Ivana Barbieri, Josip Čičak FORENSIC ACCOUNTING IN SUSTAINABILITY REPORTING: A BIBLIOMETRIC REVIEW. .	189
Dejan Jakšić, Kristina Peštović FORENSIC PROCESS MINING: A PYTHON-BASED APPROACH TO DETECTING RED FLAGS IN BUSINESS EVENT LOGS	211
Wolfgang Raeder, Jelena Demko-Rihter DESIGN-LEVEL FORENSICS: USING INTELLIGENT AGENTS TO SPOT FRAUD VULNERABILITIES IN SAP CONFIGURATION	225
Jovana Stojanović, Dejan Spasić INTERCONNECTION BETWEEN FINANCIAL DISTRESS AND THE RISK OF FRAUDULENT FINANCIAL REPORTING: EVIDENCE FROM BANKRUPT COMPANIES IN SERBIA.	237
Mario Abela ANATOMY OF A SCANDAL: DWS FROM AN INSTITUTIONAL LOGICS PERSPECTIVE. . .	257

Bobana Čegar Babić, Tajana Serdar Raković

FINANCIAL STATEMENT MANIPULATION AND CAPITAL MARKET INTEGRITY:

THE ROLE OF FORENSIC ACCOUNTING IN DETECTING EARNINGS MANAGEMENT . . . 273

Ivana Vuković, Dejan Jovanović, Mirjana Todorović

THE RELATIONSHIP BETWEEN THE ENVIRONMENTAL COMPONENT OF ESG

REPORTING AND THE FINANCIAL PERFORMANCE OF EUROPEAN COMPANIES 287

Amira Pobrić

IDENTIFYING COMPANIES WITH AN INCREASED RISK OF FRAUDULENT FINANCIAL

REPORTING USING ACCOUNTING VARIABLES. 311

Marina Vasilic

FORENSIC ACCOUNTING AND THE PROTECTION OF MINORITY SHAREHOLDERS IN

EMERGING MARKETS 331

Milan Čupić, Nikola Nikolić, Stefan Vržina

STOCK LIQUIDITY AND DIVIDEND PAYOUT OF JOINT-STOCK COMPANIES

IN THE REPUBLIC OF SERBIA 343

Snežana Miletić, Siniša Jović

ELEMENTS OF MTIC AND CAROUSEL VAT FRAUD OUTSIDE THE EU:

A FORENSIC COMPARISON. 359

Andrea Arbula Blecich, Andrija Herceg

FORENSIC ACCOUNTING IN DETECTING FINANCIAL MALFEASANCE

AND FRAUD IN NON-PROFIT ORGANIZATIONS. 383

Goran Petković

RISK ASSESSMENT OF SERVICE PROVIDERS' BUSINESS OPERATIONS –

A COMPARATIVE OVERVIEW OF DIFFERENCES IN APPROACH USING

THE EXAMPLE OF CEFTA MEMBERS. 397

Abiola Adeniyi, Kingsley Opoku Appiah, Beatrice Sarpong-Danquah

THE DIGITAL FRAUD DIAMOND: SPECIFYING THE DIGITAL DIMENSION OF CAPABILITY . . 413

Jovana Milenović, Ružica Petrović

LEGAL RISK AND FIRM MARKET VALUE: AN ANALYSIS OF THE RELATIONSHIP

BETWEEN LITIGATION AND TOBIN'S Q RATIO OF BELEX15 INDEX COMPANIES 443

Nikola Vidović

MACROECONOMIC RESILIENCE AND INSTITUTIONAL VULNERABILITY: PUBLIC SECTOR

TRANSPARENCY, FORENSIC AUDIT, AND THE MITIGATION OF SYSTEMIC FRAUD 455

Željko Vujinović

A METHODOLOGICAL FRAMEWORK FOR INTERPRETING LEGAL PROVISIONS

THAT REFER TO THE ACCOUNTING REGULATORY FRAMEWORK:

A FORENSIC ACCOUNTING APPROACH. 471

Borko Nikolić

INSTITUTIONAL MODELS OF FORENSIC ACCOUNTING EXPERTISE

IN FINANCIAL CRIME PROCEEDINGS: COMPARATIVE PERSPECTIVES

AND IMPLICATIONS FOR SERBIAN LAW 491

Biljana Simić

FROM DOCUMENTATION TO ECONOMIC SUBSTANCE:

FORENSIC INDICATORS OF TAX RISKS AND TRANSFER PRICING MANIPULATION

IN THE ERA OF ARTIFICIAL INTELLIGENCE 513

SCIENTIFIC COMMITTEE

Dejan Malinić, University of Belgrade – Faculty of Economics and Business, Serbia, Chair

Mario Abela, City University of New York - John Jay College of Criminal Justice, USA

Atanasko Atanasovski, Ss Cyril and Methodius University, Faculty of Economics, North Macedonia

Daniel Blab, Universität Regensburg - Fakultät für Wirtschaftswissenschaften, Germany

Zorica Božinovska Lazarevska, Ss Cyril and Methodius University, Faculty of Economics, North Macedonia

Josip Čičak, University of Rijeka, Faculty of Economics and Business, Croatia

Diana Dumitrescu, West University of Timisoara - Faculty of Economics and Business Administration, Romania

Dejan Jakšić, University of Novi Sad – Faculty of Economics, Serbia

Ana Lalević Filipović, University of Podgorica – Faculty of Economics, Montenegro

Igor Lončarski, University of Ljubljana – Faculty of Economics and Business, Slovenia

Dragan Mikerević, University of Banja Luka – Faculty of Economics, Bosnia and Herzegovina

Vlade Milićević, University of Belgrade – Faculty of Economics and Business, Serbia

Marija Pantelić, University of Belgrade – Faculty of Economics and Business, Serbia

Biljana Pešalj, Fontys University of Applied Sciences, Netherlands

Zabihollah Rezaee, University of Memphis, USA; Corvinus University of Budapest, Hungary

Kristina Rudžionienė, Vilnius University Kaunas Faculty - Institute of Social Sciences and Applied Informatics, Lithuania

Dejan Spasić, University of Belgrade – Faculty of Economics and Business; University of Niš – Faculty of Economics, Serbia

Aleksandar Stajković, University of Wisconsin-Madison – Wisconsin School of Business, USA

Milan Škulić, University of Belgrade – Faculty of Law, Serbia

Mirjana Todorović, University of Kragujevac – Faculty of Economics, Serbia

Tina Vuko, University of Split – Faculty of Economics, Business and Tourism, Croatia

ORGANISING COMMITTEE

Savka Vučković Milutinović, University of Belgrade – Faculty of Economics and Business, Serbia, Chair

Vukašin Kuč, University of Belgrade – Faculty of Economics and Business, Serbia

Snežana Miletić, University of Belgrade – Faculty of Economics and Business, Serbia

Tijana Ostojić, University of Belgrade – Faculty of Economics and Business, Serbia

Siniša Radić, University of Belgrade – Faculty of Economics and Business, Serbia

Mladen Stamenković, University of Belgrade – Faculty of Economics and Business, Serbia

Saša Veljković, University of Belgrade – Faculty of Economics and Business, Serbia

Milutin Živanović, University of Belgrade – Faculty of Economics and Business, Serbia

REVIEWERS

Atanasko Atanasovski, Ss Cyril and Methodius University - Faculty of Economics, North Macedonia

Vanja Bajović, University of Belgrade – Faculty of Law, Serbia

Miloš Božović, University of Belgrade – Faculty of Economics and Business, Serbia

Bobana Čegar Babić, University of Banja Luka – Faculty of Economics, Bosnia and Herzegovina

Milan Čupić, University of Kragujevac – Faculty of Economics, Serbia

Jelena Demko Rihter, University of Novi Sad – Faculty of Technical Sciences, Serbia

Ksenija Denčić-Mihajlov, University of Niš – Faculty of Economics, Serbia

Nikola Fabris, University of Belgrade – Faculty of Economics and Business, Serbia

Dejan Jakšić, University of Novi Sad – Faculty of Economics, Serbia

Dejan Jovanović, University of Kragujevac – Faculty of Economics, Serbia

Milica Kolaković Bojović, Institute of Criminological and Sociological Research, Serbia

Svetislav Kostić, University of Belgrade – Faculty of Law, Serbia

Vukašin Kuč, University of Belgrade – Faculty of Economics and Business, Serbia

Milan Lakićević, University of Podgorica – Faculty of Economics, Montenegro

Ana Lalević Filipović, University of Podgorica – Faculty of Economics, Montenegro

Dejan Malinić, University of Belgrade – Faculty of Economics and Business, Serbia

- Dragan Mikerević**, University of Banja Luka – Faculty of Economics, Bosnia and Herzegovina
- Snežana Miletić**, University of Belgrade – Faculty of Economics and Business, Serbia
- Vlade Milićević**, University of Belgrade – Faculty of Economics and Business, Serbia
- Sunčica Milutinović**, University of Novi Sad – Faculty of Economics, Serbia
- Dorđe Mitrović**, University of Belgrade – Faculty of Economics and Business, Serbia
- Marija Pantelić**, University of Belgrade – Faculty of Economics and Business, Serbia
- Jelena Perović**, University of Belgrade – Faculty of Economics and Business, Serbia
- Biljana Pešalj**, Fontys University of Applied Sciences, Netherlands
- Kristina Peštović**, University of Novi Sad – Faculty of Economics, Serbia
- Jelena Poljašević**, University of Banja Luka – Faculty of Economics, Bosnia and Herzegovina
- Siniša Radić**, University of Belgrade – Faculty of Economics and Business, Serbia
- Dejan Spasić**, University of Belgrade – Faculty of Economics and Business, Serbia,
University of Niš – Faculty of Economics, Serbia
- Vojislav Sekerez**, University of Belgrade – Faculty of Economics and Business, Serbia
- Mladen Stamenković**, University of Belgrade – Faculty of Economics and Business, Serbia
- Jelena Stanković**, University of Niš – Faculty of Economics, Serbia
- Milan Škulić**, University of Belgrade – Faculty of Law, Serbia
- Mirjana Todorović**, University of Kragujevac – Faculty of Economics, Serbia
- Savka Vučković Milutinović**, University of Belgrade – Faculty of Economics and Business
- Tina Vuko**, University of Split – Faculty of Economics, Business and Tourism, Croatia
- Bojana Vuković**, University of Novi Sad – Faculty of Economics, Serbia
- Milutin Živanović**, University of Belgrade – Faculty of Economics and Business, Serbia

FOREWORD

Fraud, in its broadest sense, is a phenomenon present to varying degrees across the world, affecting virtually all countries, regardless of their size, level of economic development, or political and social systems. It takes many forms, including corruption, the misuse of public funds, and the manipulation of business transactions, which may serve as a basis for fraudulent financial reporting. Such practices lie at the heart of major financial scandals that have shaken public confidence in institutions and attracted widespread attention in countries around the world. The greater visibility of fraud in developed countries compared with less developed ones does not necessarily indicate a lower incidence of fraud in the latter. Rather, it may reflect lower levels of transparency in fraud reporting, weaker institutional frameworks, limited effectiveness in legal proceedings, and other systemic deficiencies that hinder the detection, disclosure, and prosecution of fraudulent activities. Although these forms of fraud differ in their underlying motives, methods of execution, and the magnitude of the losses they cause, they share several fundamental characteristics: the unlawful and unethical conduct of perpetrators, the abuse of entrusted authority and responsibility, the pursuit of illicit gains at the expense of stakeholders who suffer the adverse consequences, and the resulting erosion of public trust in the institutions responsible for preventing, detecting, and sanctioning such misconduct. Of particular concern is the pervasive nature of fraud, which ranges from relatively minor schemes that often remain concealed from public scrutiny to large-scale fraudulent activities capable of destabilizing entire national economies. This concern is further compounded by the recognition that the number of undetected fraud cases is likely to far exceed the number of those uncovered and prosecuted. Some estimates suggest that only approximately one-third of fraud cases are detected, while the actual financial and broader economic and social costs may be substantially greater than those that can be explicitly identified and measured.

When examined in a broader context, the adverse effects of fraud cannot be assessed solely in terms of the direct financial losses suffered by individual businesses, investors, creditors, public institutions, and other stakeholders. This is primarily because the consequences of fraudulent activities extend beyond the immediate victims and may spread throughout the entire national economy, often in ways that are difficult to reliably quantify. Moreover, the frequency of different types of fraud does not necessarily correspond to the magnitude of the damage they cause. For example, fraudulent financial reporting occurs considerably less frequently than asset misappropriation and corruption, yet its adverse consequences can be substantially more far-reaching. The misuse of public funds, with corruption as one of its principal manifestations, undermines the efficiency of public expenditure, facilitates illicit enrichment, exacerbates social inequalities, weakens fiscal stability, reduces the quality of public services, and erodes trust in public institutions. These effects, in turn,

constrain the resources available for public investment in infrastructure, education, healthcare, culture, and other essential areas of societal development. Fraud in the private sector also generates costs that extend beyond readily identifiable financial losses. These less visible consequences include higher costs of doing business, an uneven competitive environment, flawed investment decisions, and the inefficient allocation of capital. Furthermore, the erosion of capital market efficiency, declining investor confidence, reduced foreign investment, heightened financial instability, job losses, lower tax revenues, and slower economic growth represent additional consequences of fraudulent conduct.

Clearly, fraud gives rise to economic and institutional risks that extend far beyond the substantial losses suffered by individual victims. The prevention of such misconduct is therefore an imperative. Effective fraud prevention, deterrence, and timely detection require coordinated action by regulatory and supervisory authorities, sound corporate governance, transparent financial reporting, and high-quality auditing. Particular emphasis must also be placed on the education of employees in public institutions responsible for combating fraud. Within this context, the quality education of forensic accountants deserves special attention. Fraud schemes are not infrequently designed or facilitated by highly educated professional advisers who possess relevant multidisciplinary expertise. Effectively addressing such schemes therefore requires an even higher level of knowledge, specialized skills, and investigative capacity among the public institutions responsible for countering fraud.

In light of the risks outlined above, including those associated with complex reporting practices vulnerable to fraud and based on the three-dimensional ESG framework, priority must be given to preventive measures, followed by the effective detection of fraud risks across a wide range of areas. In this regard, there is little doubt that forensic accountants must be recognized as guardians of the public interest, a role that requires substantial investment in developing the advanced competencies necessary for effective forensic accounting practice. In the ongoing effort to counter fraud, considerable expectations are placed on the application of artificial intelligence tools for forensic purposes, including machine learning (ML), natural language processing (NLP), and generative artificial intelligence (GenAI), as well as complementary technologies such as robotic process automation (RPA) and optical character recognition (OCR). Machine learning algorithms can be applied across entire datasets to detect unusual transactions, identify complex fraud schemes more rapidly, uncover subtle anomalies, and identify the effects of related-party transactions. They can also analyze extensive volumes of non-financial information by examining relevant regulations, contracts, emails, and other pertinent unstructured documentation. The application of predictive analytics can further assist investigators in identifying fraud risks at an early stage, thereby supporting timely intervention and more effective fraud prevention.

Undoubtedly, the potential contribution of artificial intelligence to fraud investigation, detection, and evidentiary processes is highly promising. However, it is equally important to recognize that fraud perpetrators will have access to the same tools. Generative artificial intelligence can be exploited to create false identities, fictitious invoices, shipping documents, and other forms of documentation designed to conceal the true economic substance of transactions, misrepresent financial information, facilitate money laundering, enable the diversion of funds, and support other fraudulent activities. By generating fictitious business transactions that convincingly replicate legitimate business activities, perpetrators may be able to circumvent existing fraud prevention and detection mechanisms. Artificial intelligence can also be used to create fraudulent investment platforms and facilitate various forms of capital market manipulation. These developments will present forensic accountants with increasingly complex challenges. In addition to dealing with established forms of fraud that have long been part of the financial and business environment, they will increasingly face highly sophisticated schemes enabled by artificial intelligence.

The use of artificial intelligence by the forensic community enhances forensic accountants' capacity to monitor and analyze complex digital trails, while, on the perpetrators' side, AI enables the creation of increasingly sophisticated fraud schemes at a significantly higher level of technical expertise. This dynamic increasingly resembles a "arms race," suggesting that the effective detection of fraud will continue to depend on advanced accounting expertise, while also requiring additional technical competencies that enable forensic accountants to understand algorithmic structures. In this regard, the academic community bears the responsibility for advancing knowledge and skills by keeping pace with emerging trends and developing curricula that integrate interdisciplinary knowledge and equip students with the competencies required to address contemporary challenges in forensic accounting. Regulatory bodies and public institutions bear an equally important responsibility to recruit and retain highly qualified forensic experts and provide adequate funding for their continuous professional development. Ultimately, forensic accountants themselves must recognize their responsibility to maintain high professional standards, commit to lifelong learning, and make the necessary investments in their own development if they are to preserve the credibility of their profession and establish themselves as genuine guardians of the public interest.

The Third Conference on Forensic Accounting, entitled "From Accounting Manipulation to Sustainability and Artificial Intelligence: New Horizons in Forensic Accounting," aims to explore the evolving landscape of financial fraud and the expanding role of forensic accounting in identifying, investigating, and addressing both traditional and emerging forms of misconduct. The objective is to bring together diverse perspectives, methodologies, and practical approaches that reflect the growing complexity of contemporary financial and business environments. The scientific program comprises 27 papers authored by 50 contributors, covering a broad spectrum of topics, including artificial intelligence and digital technologies in fraud

detection, forensic process mining, intelligent systems and digital fraud risks, fraudulent financial reporting, tax fraud and transfer pricing manipulation, capital market manipulation, ESG-related fraud and greenwashing, sustainability reporting, and the role of auditing, professional skepticism, and institutional frameworks in combating financial misconduct. Particular attention is also devoted to the legal and regulatory dimensions of forensic accounting, the protection of investors and minority shareholders, and the prevention of fraud in both the public and private sectors. The participation of contributors from 11 countries, spanning Serbia, the wider region, and the international academic community, underscores the Conference's growing international reach and its role in fostering cross-border collaboration and the exchange of knowledge and experience in addressing contemporary challenges in forensic accounting. By bringing together research on established forms of financial misconduct and emerging challenges associated with digital transformation, sustainability, and technological innovation, the Conference seeks to advance interdisciplinary dialogue and contribute to the further development of forensic accounting theory and practice.

The successful realisation of the Third Conference on Forensic Accounting would not have been possible without the engagement and support of many individuals and institutions. We are especially grateful to our keynote speakers and authors for sharing their knowledge and research, to the reviewers for their careful assessment of the submitted papers, and to the members of the Programme and Organising Committees for their contribution to the preparation and implementation of the Conference. Our appreciation also goes to the students from all three generations of the Master's programme in Forensic Accounting, whose participation added an important educational dimension to the Conference. Their involvement in the practical case study exercises on the second day provided a valuable opportunity to connect academic knowledge with practical challenges in the field. Finally, we gratefully acknowledge the support of the INL and its contribution to initiatives aimed at advancing forensic accounting and developing relevant professional capacities in the Republic of Serbia. This cooperation remains important for fostering knowledge exchange and further strengthening the field.

Editors

Douglas Cumming
School of Business, Stevens Institute
of Technology
ORCID: 0000-0003-4366-6112

FRAUD, MANIPULATION, AND REGULATORY DESIGN: WHAT EVIDENCE FROM MATURE MARKETS CAN AND CANNOT TELL SERBIA AND THE REGION

UDK 343.53:336.76(497.11)

DOI: <https://doi.org/10.54318/czf.2026.dc.001>

Abstract

This invited paper, based on a keynote lecture, reviews the evidence on financial fraud and market manipulation and asks what it implies for regulation and enforcement in Serbia and the Western Balkans. The empirical literature is built mostly on data from large markets. It shows that misconduct is common and mostly undetected, that its costs fall heavily on firm reputation, liquidity, real investment, and household trust, and that rules on the books matter far less than the surveillance and enforcement behind them. I organise the evidence around the economics of crime: the gain from misconduct, the probability of detection, and the sanction. Two parallel sections explain how market trading misconduct (spoofing, wash trading, marking the close, insider trading) and accounting misconduct (misstatement, hidden liabilities, fictitious assets) are detected and measured, illustrated with landmark cases and with evidence from Serbia and its neighbours. Separate tables summarise the trading and accounting literatures, a third table summarises landmark cases, a fourth translates the evidence into design choices for Serbia, and a fifth lists data sources for large-sample research. Three lessons transfer directly: enforcement must be credible before rules are elaborated; detection depends on employees, whistleblowers, auditors, and cooperating agencies as much as on regulators; and enforcement capacity, meaning supervisory staff, procedural tools, and workable sanctions, matters as much as new rules. Other lessons need adaptation, because thin trading, concentrated ownership, weak private enforcement, and scarce data change both the incentives to commit misconduct and the tools available to detect it. The paper closes with a research agenda built on regional datasets of enforcement actions, financial statements, trading, and detected fraud.

Key words: *fraud; market manipulation; insider trading; enforcement; market surveillance; Serbia; Western Balkans*

JEL Classification: *G14, G18, G38, K22, K42, M42*

Introduction

Financial fraud and market manipulation are not rare events at the margins of capital markets. Dyck, Morse, and Zingales (2024) estimate that only about one third of corporate frauds are detected in normal times, that about 10% of large publicly traded firms in the United States (US) are committing securities fraud in any given year, and that fraud destroys about 1.6% of equity value each year, roughly USD 830 billion in 2021. In secondary markets, Comerton-Forde and Putniņš (2014) estimate that about 1% of closing prices on the exchanges they study are manipulated, and that only a small fraction of these cases is detected and prosecuted. Practitioner evidence agrees: the Association of Certified Fraud Examiners (ACFE, 2024) analyses 1,921 occupational fraud cases from 138 countries and territories and reports that 43% were detected through tips, more than three times as many as through any other method.

The costs extend well beyond the amounts stolen or mispriced. The reputational loss to firms caught misrepresenting their accounts is more than 7.5 times the sum of all legal and regulatory penalties (Karpoff, Lee, & Martin, 2008b), and 93% of the managers held responsible lose their jobs (Karpoff, Lee, & Martin, 2008a). Misconduct also reduces liquidity, distorts real decisions such as innovation and acquisitions, and erodes the trust that brings households into equity markets (Giannetti & Wang, 2016).

Most of this evidence comes from large and liquid markets. The question for regulators, prosecutors, police and tax investigators, auditors, and forensic accountants in Serbia and the wider Western Balkans is which findings travel to markets that are much smaller and thinner. In 2025 the Belgrade Stock Exchange recorded total turnover of about EUR 210 million, of which about EUR 197 million was in bonds and only about EUR 13 million in shares (Belgrade Stock Exchange, 2026). At the same time, Serbia's Law on Capital Market, adopted in 2021 and with most provisions applied since January 2023, brought domestic rules into line with EU legislation on financial instruments, prospectuses, transparency, and market abuse (Law on Capital Market, 2021), and further alignment with the EU Market Abuse Regulation (MAR) is under way as part of the accession process (Policy and Legal Advice Centre [PLAC IV], 2025). Rules are converging with those of mature markets; the conditions in which they operate are not.

This paper reviews the evidence on fraud, manipulation, and enforcement with that gap in mind. It draws on my own work with co-authors on exchange trading rules, market surveillance, and enforcement, updates the synthesis in Cumming, Dannhauser, and Johan (2015), and brings in the growing body of research from Serbia and neighbouring markets. Section 1 sets out a taxonomy of misconduct. Section 2 presents the economic framework and the evidence on why misconduct occurs. Sections 3 and 4 explain, in parallel, how market trading misconduct and accounting misconduct are detected and measured, with landmark cases and regional evidence. Section 5 reviews the consequences of misconduct and Section 6 the

evidence on regulatory design. Section 7 turns to Serbia and the region: the current setting, the lessons that transfer and those that need adaptation, and a research agenda. Figure 1 summarises the framework, Tables 1 and 2 summarise the trading and accounting literatures, Table 3 summarises landmark cases, Table 4 translates the evidence into design choices, and Table 5 lists data sources for large-sample research in Serbia and the region.

1. Forms of fraud and manipulation: a working taxonomy

Misconduct in financial markets falls into two broad families that are often studied separately but overlap in practice. The first is corporate fraud: financial statement fraud, asset misappropriation, and corruption, the three categories practitioners use for occupational fraud (ACFE, 2024), together with false or misleading disclosure to investors. The second is trading misconduct: insider trading and market manipulation. Allen and Gale (1992) show that even an uninformed trader can profit from manipulation as long as other investors attach a positive probability to the manipulator being informed. Manipulation is therefore not self-correcting, and some of it survives in equilibrium unless it is deterred.

Exchange rules are far more detailed than a general prohibition of manipulation. Cumming, Johan, and Li (2011) code the trading rules of 42 exchanges into indices covering market manipulation (including price manipulation such as marking the close and ramping, volume manipulation such as wash trades, and spoofing), insider trading (including front-running of client orders), and broker–agency conflicts. They show that differences in these rules, across markets and over time, significantly affect liquidity. Cumming, Dannhauser, et al. (2015) build on this taxonomy to relate the many forms of misconduct to agency conflicts between managers, brokers, and investors. Alexander and Cumming (2020) bring together research on the detection, prevention, and regulation of these practices, from insider trading, front-running, bribery, and false disclosure to wash trades, spoofing, and manipulation on cryptocurrency exchanges.

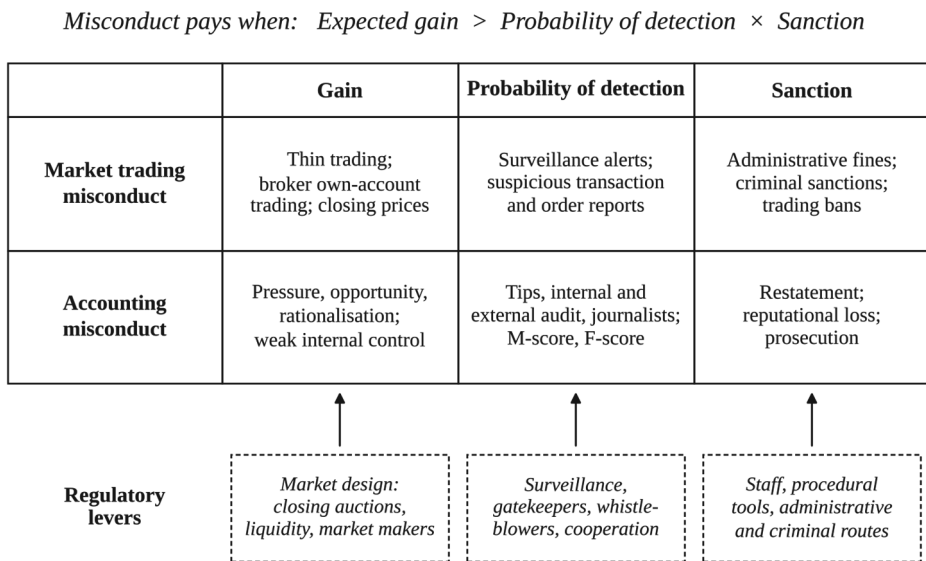
The boundary between the two families is porous, which matters for investigators. An accounting misstatement can be accompanied by insider trading before its disclosure, and manipulation of a closing price can alter the outcome and terms of a takeover (Cumming, Ji, Johan, & Tarsalewska, 2020). Misconduct also migrates to new venues: fraud in reward-based crowdfunding (Cumming, Hornuf, Karami, & Schweizer, 2023) hacks of cryptocurrency exchanges, wallets, and blockchains (Corbet, Cumming, Lucey, Peat, & Vigne, 2020), and crypto-asset trading, where wash trading accounts for more than 70% of reported volume on unregulated exchanges (Cong, Li, Tang, & Yang, 2023) and openly announced pump-and-dump schemes produce price distortions of 65% on average (Dhawan & Putniņš, 2023). This is not remote from Serbia, whose Law on Digital Assets has applied since 2021, with the National Bank of Serbia supervising virtual currencies and the Securities Commission supervising digital tokens (Law on Digital Assets, 2020). Forensic

accountants and market surveillance staff therefore often work on the same case from different ends.

2. Why misconduct occurs and how it is detected

The most useful organising framework remains the economics of crime (Becker, 1968). A rational wrongdoer commits misconduct when the expected gain exceeds the expected cost, which is the probability of detection multiplied by the sanction. Each regulatory lever works through one of these terms: market design changes the gain, surveillance and gatekeepers change the probability of detection, and penalties and reputational loss change the sanction. Figure 1 applies the framework to the two families of misconduct reviewed in this paper. For accounting misconduct, the gain side is often described through the conditions identified by Cressey (1953) in his study of embezzlement: a pressure or need that cannot be shared, a perceived opportunity, and a rationalisation.

Figure 1: The economics of misconduct applied to market trading misconduct and accounting misconduct: what drives the expected gain, the probability of detection, and the sanction, and which regulatory levers act on each



Note: Following Becker (1968), misconduct pays when the expected gain exceeds the probability of detection multiplied by the sanction. Rows show how each term arises for the two families of misconduct; the bottom row shows the regulatory levers that act on each term (arrows). For accounting misconduct, the gain side reflects the pressure, opportunity, and rationalisation identified by Cressey (1953). M-score and F-score are statistical screens for earnings manipulation and misstatement.

Source: Author's elaboration based on Becker (1968) and Cressey (1953)

2.1. Opportunity and the gains from misconduct

Gains are largest where prices are easy to move and where the wrongdoer controls the flow of orders or information. Aggarwal and Wu (2006) study US enforcement cases and find that manipulators are typically plausibly informed parties such as insiders and brokers, that prices rise during the manipulation and fall afterwards, and that manipulation raises volatility, liquidity, and returns while it lasts; most of their cases involve small, illiquid stocks traded off the main exchanges. The most direct evidence from a smaller emerging market comes from Khwaja and Mian (2005), who show that brokers in Pakistan trading on their own account earned annual returns 50 to 90 percentage points higher than outside investors through trade-based pump-and-dump schemes, and that manipulation rents could account for almost half of total broker earnings.

Opportunity also changes with circumstances. Comerton-Forde and Putniņš (2014) find that regulatory budgets strongly affect both the incidence and the detection of manipulation. Manipulation becomes more frequent and more severe during natural disasters (Aker, Cumming, & Ji, 2023), while internal data from a bank show that working from home lowered the likelihood of securities misconduct among traders (Cumming, Firth, Gathergood, & Stewart, 2023), a reminder that opportunity depends on the working environment as well as on formal controls.

2.2. Governance, gatekeepers, and detection

Internal governance shapes both the incidence of fraud and the market's response to it. Greater board gender diversity is associated with less frequent fraud among Chinese listed firms and with a milder market reaction when fraud occurs (Cumming, Leung, & Rui, 2015). External monitors matter as well: analyst coverage deters fraud, especially in non-state firms that rely more on capital markets (Chen, Cumming, Hou, & Lee, 2016), and modified audit opinions moderate the link between low executive integrity and subsequent enforcement (Chen, Cumming, Hou, & Lee, 2013).

The single most important finding for practitioners is that detection “takes a village”. Dyck, Morse, and Zingales (2010) show that fraud in large US firms is uncovered not mainly by the standard governance actors, namely investors, the Securities and Exchange Commission (SEC), and auditors, but by employees, the media, and industry regulators. Short sellers build positions in the months before misrepresentation is revealed and speed its discovery (Karpoff & Lou, 2010). In the practitioner evidence, just over half of all tips (52%) come from employees (ACFE, 2024). Whistleblowers also improve outcomes once a case is opened: their involvement is associated with higher penalties, longer prison sentences for culpable executives, and faster enforcement (Call, Martin, Sharp, & Wilde, 2018), and larger financial rewards under the US False Claims Act increased whistleblowing, although whistleblowers bore lasting income losses (Dey, Heese, & Pérez-Cavazos, 2021). Auditors matter less as detectors than as a deterrent: fraudulent reporting is less likely

when the auditor is a Big Five firm (Lennox & Pittman, 2010). Because most frauds are never detected (Dyck et al., 2024), the channels through which information reaches investigators are as important as the investigators themselves.

3. Detecting market trading misconduct

3.1. What surveillance looks for

Market abuse leaves traces in order and trade data, and regulators describe the patterns they look for. The EU Market Abuse Regulation lists indicators of manipulation, including the extent to which orders or transactions represent a significant proportion of daily volume, whether transactions lead to no change in beneficial ownership, and the extent to which orders change the best bid or offer and are removed before they are executed (Regulation (EU) No 596/2014, 2014, Annex I). The implementing rules describe the main practices in more detail (Commission Delegated Regulation (EU) 2016/522, 2016, Annex II). Spoofing and layering mean submitting large orders, often away from the best price, on one side of the order book in order to execute a trade on the other side, and then cancelling the orders that were never meant to trade. Wash trades are transactions that involve no change in beneficial interest or market risk, so that volume or prices are painted without any real transfer. Marking the close means trading at the reference time, such as the close or the settlement, to move the reference price on which valuations, margin calls, and contracts depend. Insider trading shows up as abnormal trading and price movements before price-sensitive announcements. The scale of the task is large: the United Kingdom's Financial Conduct Authority reports processing around 150 million order book messages a day and running surveillance algorithms for spoofing and layering (Financial Conduct Authority [FCA], 2021).

Much of this surveillance is now carried out by computer algorithms. The SMARTS system, developed by Michael Aitken at the University of New South Wales, flags unusual price or volume movements before information is published, and by 2010 it was used by more than 40 exchanges and regulators (University of New South Wales [UNSW], 2010). Nasdaq OMX acquired SMARTS in 2010 (NASDAQ OMX Group, 2010), and Nasdaq now reports that its market surveillance platform is used by more than 50 marketplaces and 20 regulators, with more than 70 pre-configured alerts and machine-learning tools to prioritise cases (Nasdaq, n.d.). Such alerts are also a research tool. Cumming, Ji, Peter, and Tarsalewska (2020) use SMARTS alerts to identify suspected end-of-day manipulation: a closing price change more than four standard deviations from its mean over the previous 100 trading days, at least half of which reverses at the next day's open. Because alerts do not depend on whether a case is prosecuted, they measure suspected misconduct independently of enforcement choices, and Akter et al. (2023) use similar alerts for US stocks.

The metrics behind such alerts can be stated simply, and they become evidence in court. In the first US criminal conviction for spoofing, only 0.57% of the defendant's large orders stayed on the market for more than one second, against 65% for other

high-frequency traders; his large orders were filled 0.08% of the time on the main exchange against 35.61% for his small orders; and his order-to-trade ratio was 1,592%, against 91% to 264% for other participants (United States v. Coscia, 2017). Each of these numbers can be computed from exchange data for any account, which is why order-level data and the staff to analyse them are the foundation of detection.

3.2. Landmark cases

Table 3 summarises landmark cases. In spoofing, a London-based trader used a layering algorithm that kept four to six large sell orders in the E-mini S&P 500 futures order book and cancelled most of them; United States authorities charged that the scheme contributed to the 6 May 2010 flash crash (Commodity Futures Trading Commission [CFTC], 2015). JPMorgan paid more than USD 920 million in 2020 to resolve spoofing in precious metals and US Treasury futures that ran from 2008 to 2016 (U.S. Department of Justice [DOJ], 2020). In marking the close, the CFTC charged Optiver with attempting to manipulate energy futures settlement prices in 2007, citing recorded telephone calls in which traders spoke of being able to “bully the market” (CFTC, 2008); the trading had been detected by the exchange's own surveillance, and the firm settled for USD 14 million (CFTC, 2012). Benchmark manipulation can be as damaging as price manipulation: Barclays was penalised by United States and British authorities in 2012 for submissions to the London and Euro interbank offered rates (LIBOR and EURIBOR) that suited traders' positions, with the British regulator relying on internal emails and messages (Financial Services Authority [FSA], 2012). In crypto-asset markets, the Federal Bureau of Investigation created its own token to catch market makers offering wash trading, leading to charges against 18 individuals and entities in 2024 (DOJ, 2024).

Three features of these cases matter for smaller markets. First, detection came from data analysis and from communications records as much as from tips. Second, the conduct often ran for years before it was sanctioned, so the deterrent effect depends on cases that are seen to end in penalties. Third, the same techniques that were used on the largest futures markets, such as order-to-trade ratios, cancellation rates, and closing-price dislocations, can be applied to the much smaller order books of the Belgrade Stock Exchange.

Table 1: Selected empirical evidence on market trading misconduct (manipulation, insider trading, spoofing, and wash trading): prevalence, detection, regulation, and evidence from Serbia and nearby markets

Study (authors, year)	Market, sample, and period	Main finding
<i>Prevalence and mechanics</i>		
Comerton-Forde & Putniņš (2014)	NYSE, AMEX, TSX, TSX-V, 1997–2008	About 1% of closing prices are manipulated; only a small fraction is detected and prosecuted.
Aggarwal & Wu (2006)	US SEC manipulation cases, 1990–2001	Manipulators are typically informed parties (insiders, brokers); prices rise during manipulation and fall afterwards.
Khwaja & Mian (2005)	Pakistan, broker-level trading data	Brokers trading on own account earn 50–90 percentage points higher annual returns through pump-and-dump schemes.
Cong, Li, Tang, & Yang (2023)	29 crypto exchanges	Wash trading exceeds 70% of reported volume on unregulated exchanges.
Dhawan & Putniņš (2023)	355 crypto pump-and-dumps	Openly announced pumps distort prices by 65% on average.
<i>Detection and surveillance</i>		
Cumming & Johan (2008)	Surveillance in 25 jurisdictions	Exchanges as self-regulators run wider single-market surveillance; information sharing strengthens cross-market surveillance.
Cumming, Ji, Peter, & Tarsalewska (2020)	SMARTS end-of-day alerts, nine countries	Suspected manipulation is associated with less innovation, more so where intellectual property rights are weak.
Aitken, Cumming, & Zhan (2015a)	22 exchanges, 2003–2011	Rules and surveillance reduce suspected insider trading cases but raise the profit per case.
Del Guercio, Odders-White, & Ready (2017)	SEC insider trading cases, 2003–2011	Greater enforcement intensity goes with smaller pre-announcement run-ups and narrower spreads.
<i>Rules, enforcement, and market design</i>		
Bhattacharya & Daouk (2002)	103 countries	87 had insider trading laws, 38 had prosecuted; the cost of equity falls after the first prosecution, not the law.
Cumming, Johan, & Li (2011); Cumming & Johan (2019)	42 exchanges; review of EU reform studies	More detailed trading rules improve liquidity; dating reforms by implementation or by signing gives the same gains.

Study (authors, year)	Market, sample, and period	Main finding
Christensen, Hail, & Leuz (2016)	EU Market Abuse and Transparency Directives	Replicates the liquidity gains in Cumming et al. (2011) with signing dates; effects stronger with strict implementation.
Cumming, Groh, & Johan (2018)	EU countries under harmonised rules	Enforcement intensity is the most robust predictor of market abuse detection.
Aitken, Cumming, & Zhan (2015b); Hillion & Suominen (2004)	International exchanges; Paris Bourse	High-frequency trading and closing call auctions reduce closing price manipulation.
<i>Evidence from Serbia and nearby markets</i>		
Kostić & Jelisavac Trošić (2020)	Serbian law and enforcement practice	Market abuse offences exist on paper but have in practice appeared impossible to prove in court.
Živković & Minović (2010)	Belgrade Stock Exchange, 2005–2009	Liquidity is low and persistent; swings in foreign participation drive illiquidity.
Kuveždić & Dedi (2022)	Zagreb insider notifications, 2010–2021	Insider purchases earn significant abnormal returns; sales do not.
Alexakis, Pappas, & Skarmeeas (2021)	Athens Stock Exchange tick data	A closing call auction reduced but did not eliminate closing price manipulation.
Matek & Galić (2024)	Zagreb and Ljubljana, 19 stocks, 2010–2022	Designated market makers narrow bid–ask spreads; effects on turnover are inconclusive.

Note: The table summarises selected studies of misconduct in trading on securities, futures, and crypto-asset markets, grouped by theme. Findings are paraphrased from each study's abstract or text and are not exhaustive. Abbreviations: AMEX = American Stock Exchange; NYSE = New York Stock Exchange; TSX and TSX-V = Toronto Stock Exchange and TSX Venture Exchange; SEC = US Securities and Exchange Commission; SMARTS = market surveillance system, now part of Nasdaq; EU = European Union. Full references are in the reference list.

Source: Author's compilation from the cited studies

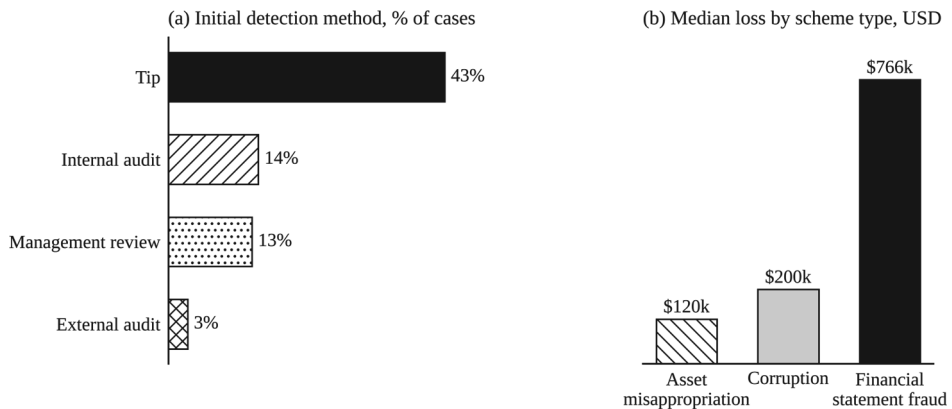
4. Detecting and measuring accounting misconduct

4.1. Schemes and costs

Accounting misconduct is less frequent than other occupational fraud but far more costly. In the practitioner evidence, financial statement fraud accounts for 5% of cases but has a median loss of USD 766,000, against USD 120,000 for asset misappropriation and USD 200,000 for corruption, and it runs for a median of 18 months before detection (ACFE, 2024; Figure 2). The schemes are familiar to forensic accountants: overstated or fictitious revenue, expenses capitalised as assets,

liabilities kept off the balance sheet, and assets, including cash, that do not exist. A survey of auditors in Serbia, Croatia, North Macedonia, and Bosnia and Herzegovina finds that Serbian auditors encounter unfounded revenue overstatement more often than auditors in the other countries and report the highest scores for inventory manipulation, and that wider legal powers for auditors would reduce the limits on fraud detection (Dimitrijević, Jovković, & Milutinović, 2021).

Figure 2: How occupational fraud is detected and what each scheme costs: (a) initial detection method and (b) median loss by scheme type, global sample of 1,921 cases from 138 countries and territories



Note: Global data on occupational fraud cases investigated by certified fraud examiners between January 2022 and September 2023, not data for a single country. The largest group of cases came from the United States and Canada (623 cases); 66 cases came from Eastern Europe and Western/Central Asia. Panel (a) shows the share of cases first detected by each method for the three most common methods and, for comparison, external audit; the methods omitted include document examination (6%), account reconciliation (5%), and detection by accident (5%), so shares do not sum to 100%. Panel (b) shows the median loss per case in thousands of US dollars for the three scheme types; a case can involve more than one scheme. Financial statement fraud accounts for 5% of cases but has the largest median loss.

Source: Author's presentation of data from the Association of Certified Fraud Examiners (ACFE, 2024)

4.2. How accounting misconduct is measured

Researchers measure accounting misconduct in two ways. The first uses detected events: restatements, regulators' enforcement releases, and securities class actions. These data are informative but incomplete, and even the US databases most widely used in research miss relevant events, record late revelation dates, and include events that are not fraud (Karpoff, Koester, Lee, & Martin, 2017). Because only about a third of frauds are detected (Dyck et al., 2024), detected events understate prevalence. The second approach estimates the discretionary part of reported earnings. Jones (1991) models normal accruals as a function of revenue growth and

fixed assets, and the modified Jones model of Dechow, Sloan, and Sweeney (1995) adjusts revenue growth for the change in receivables, so that credit sales cannot be used to inflate earnings without being flagged. Accruals above what the model predicts are treated as discretionary. Using an aggregate measure built from several indicators of accrual-based earnings management, Miletić and Vučković Milutinović (2020) find, for 215 large Serbian companies from 2009 to 2017, that such earnings management is widespread.

4.3. Detection tools

For forensic accountants, tax inspectors, and auditors, accounting-based screens are a practical starting point. Beneish (1999) develops a model, now known as the M-score, that combines eight ratios, among them the growth of receivables relative to sales, gross margin, asset quality, sales growth, depreciation, overhead costs, leverage, and total accruals, and that identifies about half of the companies involved in earnings manipulation before public discovery. Dechow, Ge, Larson, and Sloan (2011) develop the F-score, a scaled probability of material misstatement built from financial statement variables such as accruals and performance, as a red flag for further review. Bao, Ke, Li, Yu, and Zhang (2020) show that a machine-learning model using raw accounting numbers outperforms both the ratio-based logistic model of Dechow et al. (2011) and an earlier support-vector-machine model. Amiram et al. (2018) review this literature across law, accounting, and finance. Such screens flag firms for closer examination rather than prove fraud, and they are most useful where full populations of financial statements are available, as they are in Serbia through the Business Registers Agency.

Two further tools are widely used in practice. Benford's law describes the distribution of first digits in many naturally occurring data sets, in which the digit 1 appears about 30% of the time and the digit 9 less than 5% of the time (Benford, 1938); departures from this pattern in ledgers, invoices, or expense claims can signal invented numbers (Nigrini, 2012). And proactive data monitoring pays: organisations that use it have median fraud losses about half as large, and detect fraud in about half the time, as those that do not (ACFE, 2024). These tools have already been applied in the region. Dimitrijević, Obradović, and Milutinović (2018) apply the Beneish model to Serbian companies and find that the risk of fraudulent reporting is not negligible, with borrowing a key factor; Spasić and Stojanović (2026) classify 20.83% of 48 Serbian companies in which bankruptcy proceedings were initiated between 2017 and 2024 as potential manipulators; Petrović, Novović, and Šoškić (2025) apply Benford's law to Serbian food companies; and Isaković-Kaplan, Demirović, and Proho (2021) find that income statement items in Bosnia and Herzegovina generally follow Benford's law, while some items warrant further forensic investigation.

4.4. Landmark cases

The landmark accounting cases in Table 3 show how varied the detection channels are. At Enron, the company restated four years of results in November 2001 because special purpose entities should have been consolidated, and it filed for bankruptcy weeks later (Enron Corp., 2001). At WorldCom, internal auditors found that line costs had been capitalised as assets, leading to an announced overstatement of more than USD 3.8 billion in 2002 (Congressional Research Service, 2002). At Parmalat, assets were overstated by at least EUR 3.95 billion through a bank account that did not exist, supported by a forged confirmation (U.S. Securities and Exchange Commission [SEC], 2003). At Satyam, the founder confessed in 2009, and the SEC later charged that revenue, income, and cash had been overstated by more than USD 1 billion using more than 6,000 fake invoices (SEC, 2011a). At Wirecard, journalists and short sellers raised concerns before the auditor could not confirm EUR 1.9 billion of cash held in trust accounts in June 2020 (Wirecard AG, 2020). At Luckin Coffee, after an anonymous report published by a short seller had alleged inflated sales, the misconduct was discovered in the course of the annual external audit, an internal investigation followed, and the SEC charged that more than USD 300 million of sales had been fabricated (SEC, 2020). Three of these cases, Parmalat, Satyam, and Wirecard, turned on cash that did not exist, which makes independent confirmation of cash and bank balances one of the most valuable audit procedures.

The region has its own examples. After Agrokor entered extraordinary administration, the re-audited accounts showed undisclosed liabilities of 3.9 billion Croatian kuna (HRK) and turned the 2015 result from a reported profit of HRK 1.2 billion into a loss of HRK 3.6 billion (Agrokor, 2017). In Serbia, the National Bank of Serbia closed four banks between 2012 and 2014, citing inadequate management and supervision (National Bank of Serbia, n.d.). In the case of Agrobanka, prosecutors alleged loans to networked companies without adequate collateral, and bank officials were arrested in 2012 (Organized Crime and Corruption Reporting Project [OCCRP], 2012); in 2021 the Higher Court in Belgrade convicted ten defendants at first instance (Danas, 2021). These cases concern lending rather than published financial statements, but they show the same pattern of related-party exposure and weak controls that accounting screens are designed to flag.

Table 2: Selected empirical evidence on accounting misconduct (financial statement fraud and earnings management): prevalence and cost, measurement and detection tools, detection channels and gatekeepers, and evidence from Serbia and nearby markets

Study (authors, year)	Market, sample, and period	Main finding
<i>Prevalence and cost</i>		
Dyck, Morse, & Zingales (2024)	Large US public firms	About one third of frauds are detected; about 10% of large firms commit securities fraud each year; 1.6% of equity value destroyed per year.
Karpoff, Lee, & Martin (2008a, 2008b)	US enforcement actions, 1978–2002 (firms) and 1978–2006 (managers)	Reputational loss exceeds 7.5 times legal penalties; 93% of responsible managers lose their jobs.
ACFE (2024)	1,921 occupational fraud cases, 138 countries and territories	Financial statement fraud: 5% of cases, median loss USD 766,000, median duration 18 months.
Giannetti & Wang (2016)	US households by state	Fraud revelation reduces stock market participation, including holdings of non-fraudulent firms.
<i>Measurement and detection tools</i>		
Jones (1991); Dechow, Sloan, & Sweeney (1995)	US firms	Discretionary accruals estimated from revenue growth and fixed assets; the modified model adjusts for receivables.
Beneish (1999)	US earnings manipulators	Eight-ratio M-score identifies about half of manipulators before public discovery.
Dechow, Ge, Larson, & Sloan (2011)	SEC enforcement releases	F-score: a scaled probability of material misstatement used as a red flag.
Bao, Ke, Li, Yu, & Zhang (2020)	US accounting fraud cases	Machine learning on raw accounting numbers outperforms ratio-based models.
Karpoff, Koester, Lee, & Martin (2017)	Four major US misconduct databases	Databases miss events, record late dates, and include non-fraud events.
<i>Detection channels and gatekeepers</i>		
Dyck, Morse, & Zingales (2010)	Frauds in large US firms, 1996–2004	Employees, media, and industry regulators detect more fraud than investors, the SEC, or auditors.
Call, Martin, Sharp, & Wilde (2018)	US financial misrepresentation enforcement actions	Whistleblower involvement is associated with higher penalties, longer sentences, and faster enforcement.
Lennox & Pittman (2010)	US public companies, 1981–2001	Fraudulent reporting is less likely with a Big Five auditor.

Study (authors, year)	Market, sample, and period	Main finding
Chen, Cumming, Hou, & Lee (2013, 2016); Cumming, Leung, & Rui (2015)	Chinese listed firms	Audit opinions, analyst coverage, and board gender diversity are associated with less fraud.
<i>Evidence from Serbia and nearby markets</i>		
Miletić & Vučković Milutinović (2020)	215 large Serbian companies, 2009–2017	Accrual-based earnings management is widespread.
Dimitrijević, Obradović, & Milutinović (2018)	Serbian companies, Beneish model	Fraud risk is not negligible; borrowing is a key factor.
Spasić & Stojanović (2026)	48 Serbian companies entering bankruptcy, 2017–2024	20.83% classified as potential manipulators by the M-score.
Dimitrijević, Jovković, & Milutinović (2021)	Survey of 51 auditors in four Western Balkan countries	Revenue overstatement and inventory manipulation more frequent in Serbia than in the other countries; auditors' powers limited.
Isaković-Kaplan, Demirović, & Proho (2021)	Bosnia and Herzegovina income statements	Items generally follow Benford's law; some warrant forensic investigation.

Note: The table summarises selected studies of misstated financial reporting and earnings management, grouped by theme. Findings are paraphrased and are not exhaustive. Discretionary accruals are the part of reported earnings not explained by revenue growth and fixed assets. The M-score (Beneish) and F-score (Dechow et al.) are statistical screens that rank firms by the likelihood of manipulation or misstatement; they flag firms for review and do not prove fraud. Benford's law describes the expected frequency of first digits in naturally occurring numbers. Abbreviations: ACFE = Association of Certified Fraud Examiners; SEC = US Securities and Exchange Commission; Big Five = the five largest international audit firms before 2002. Full references are in the reference list.

Source: Author's compilation from the cited studies

Table 3: Landmark cases of market trading misconduct and accounting misconduct: the scheme, how it was detected, and the outcome

Case, market, and years (source)	Scheme	How it was detected	Outcome
<i>Market trading misconduct</i>			
Sarao, US futures, 2010–2015 (CFTC, 2015)	Layering algorithm in E-mini S&P 500 futures	CFTC investigation; press reports credit a whistleblower's analysis	Guilty plea 2016; charged as contributing to the 2010 flash crash
Coscia, US and UK futures, 2011 (United States v. Coscia, 2017)	Large orders cancelled within fractions of a second	Order duration, fill rates, order-to-trade ratio	First US criminal spoofing conviction (2015); three years in prison
JPMorgan, 2008–2016 (DOJ, 2020)	Spoofing in precious metals and Treasury futures	Investigations by DOJ, CFTC, and SEC	More than USD 920 million (2020); traders convicted (2022)
Optiver, US energy futures, 2007 (CFTC, 2008, 2012)	Banging the close at settlement	Exchange (NYMEX) surveillance; recorded telephone calls	USD 14 million settlement (2012); trading bans
Barclays, LIBOR and EURIBOR, 2005–2009 (FSA, 2012)	Benchmark submissions to suit trading positions	Press scrutiny; internal emails and messages	US and UK penalties (2012), including GBP 59.5 million from the FSA
Crypto market makers, 2024 (DOJ, 2024)	Wash trading of crypto tokens	FBI-created token to catch offers of wash trading	18 individuals and entities charged
<i>Accounting misconduct</i>			
Enron, US, 2001 (Enron Corp., 2001)	Special purpose entities kept off the balance sheet	Internal warning; restatement of 1997–2000	Bankruptcy December 2001
WorldCom, US, 2002 (Congressional Research Service, 2002)	Line costs capitalised as assets; over USD 3.8 billion announced	Internal audit	Bankruptcy July 2002; criminal convictions
Parmalat, Italy, 2003 (SEC, 2003)	Non-existent bank account of at least EUR 3.95 billion; forged confirmation	Auditors' questions about accounts	Insolvency; criminal convictions

Case, market, and years (source)	Scheme	How it was detected	Outcome
Satyam, India, 2009 (SEC, 2011a, 2011b)	Fictitious cash and revenue of more than USD 1 billion (SEC charges)	Founder's confession letter	SEC penalty USD 10 million; auditor affiliates USD 6 million
Wirecard, Germany, 2020 (Wirecard AG, 2020)	EUR 1.9 billion of trust-account cash not confirmed	Journalists and short sellers; auditor's cash confirmation	Insolvency June 2020; criminal trial ongoing, no verdict as of October 2026
Luckin Coffee, China and US, 2019–2020 (SEC, 2020)	More than USD 300 million of fabricated sales (SEC charges)	Annual external audit, then internal investigation (after an anonymous short-seller report)	SEC penalty USD 180 million
Agrokor, Croatia, 2017 (Agrokor, 2017)	Undisclosed liabilities of HRK 3.9 billion	Group crisis; re-audit under extraordinary administration	Special administration law; indictment confirmed 2026, no verdict as of October 2026
Agrobanka, Serbia, 2012 (OCCRP, 2012; Danas, 2021)	Alleged loans to networked companies without adequate collateral	National Bank of Serbia administration	Licence revoked 2012; arrests 2012; ten defendants convicted at first instance (2021)

Note: Cases were selected to illustrate the main schemes and detection channels discussed in Sections 3 and 4. Amounts and outcomes are as reported by the cited authorities, courts, or companies; outcomes are stated as of October 2026, and alleged conduct in cases without a final judgment is described as alleged. Abbreviations: CFTC = US Commodity Futures Trading Commission; DOJ = US Department of Justice; FBI = US Federal Bureau of Investigation; FSA = UK Financial Services Authority; SEC = US Securities and Exchange Commission; OCCRP = Organized Crime and Corruption Reporting Project; LIBOR and EURIBOR = London and Euro interbank offered rates; HRK = Croatian kuna; E-mini S&P 500 = stock index futures contract.

Source: Author's compilation from the cited regulatory, court, and company records

5. Consequences of fraud and manipulation

The consequences fall on the firm, on the market, and on the real economy. For the firm, reputational losses dwarf legal penalties (Karpoff et al., 2008b), a pattern reviewed for non-compliance with market regulation more broadly by Cumming, Dannhauser, and Johan (2021). For the market, misconduct and the rules against it affect liquidity (Cumming et al., 2011), and manipulation distorts the prices on which contracts, margin calls, and fund valuations depend.

The real effects are less visible but no less important. Suspected end-of-day manipulation across nine countries is associated with less corporate innovation, particularly where intellectual property rights are weak (Cumming, Ji, Peter, et al., 2020). Manipulation of a target's closing price increases the probability that a takeover is withdrawn and reduces the premium paid, while more detailed exchange trading rules lower the probability of withdrawal and mitigate the effect of manipulation on it (Cumming, Ji, Johan, et al., 2020). Fraud also has spillovers on honest firms: crowdfunding projects launched around the public detection of a fraud case are less likely to succeed and raise less money (Cumming, Hornuf, et al., 2023), and households reduce their stock holdings after fraud is revealed in their state, including holdings of firms that committed no fraud (Giannetti & Wang, 2016). In a small market where investors judge the market as a whole rather than individual issuers, one would expect such spillovers to be larger, although this has not yet been tested.

6. Regulatory design: what the evidence shows

6.1. Laws on the books versus enforcement

The clearest result in this literature is that rules without enforcement achieve little. Bhattacharya and Daouk (2002) show that of 103 countries with stock markets, 87 had insider trading laws but only 38 had ever prosecuted a case, and that the cost of equity did not change when a law was introduced but fell significantly after the first prosecution. Bhattacharya and Daouk (2009) go further: in some countries that enact an insider trading law but do not enforce it, the cost of equity actually rises, so that no law may be better than a good law that is ignored. Coffee (2007) argues that enforcement intensity, rather than formal law or legal origin, explains much of the difference between the US and other markets, and Del Guercio, Odders-White, and Ready (2017) find that greater SEC enforcement intensity is associated with smaller price run-ups before takeover and earnings announcements and with narrower bid–ask spreads. The most relevant evidence for transition economies comes from Glaeser, Johnson, and Shleifer (2001): in the 1990s, Poland's strictly enforced securities regulation was associated with market development, while the Czech Republic's reliance on minimal regulation and the courts was associated with stagnation.

Which enforcement matters is debated. La Porta, Lopez-de-Silanes, and Shleifer (2006) find that mandatory disclosure and private enforcement through liability rules benefit stock markets, with little evidence for public enforcement. Jackson and Roe (2009) measure public enforcement by the budgets and staff of regulators and reverse that result: public enforcement is as important as disclosure and more important than private liability rules. Firm-level data help reconcile the two views. Using 45,220 firms in 46 countries, Cumming, Knill, and Richardson (2015) show that public enforcement facilitates security issuance by small firms, while private enforcement benefits mainly large firms.

6.2. Harmonised rules, divergent enforcement

The EU experience is directly relevant to accession candidates. Cumming et al. (2011) show that changes in exchange trading rules, including those that implemented the EU market abuse rules, improved liquidity, dating each change by when it was implemented together with surveillance. Christensen, Hail, and Leuz (2016) replicate this analysis for the EU Market Abuse and Transparency Directives with different dates, those on which the rules were legally signed, and also find that liquidity increased; they add that the effects were stronger in countries with stricter implementation and traditionally more stringent regulation, so that countries diverged more after harmonising their rules. Cumming and Johan (2019) review the two studies and show that both dating approaches lead to the same conclusion: regulatory change improves liquidity. Cumming, Groh, and Johan (2018) show that enforcement intensity differs widely across Europe under the same rules and is the most robust predictor of market abuse detection, with the number of supervisory staff, formal cooperation between authorities, and custodial sentences all mattering. Detected fraud also varies across countries and across exchanges within a country, and the evidence points to weaker enforcement outside the US (Cumming & Johan, 2013).

6.3. Sanctions: criminal and administrative routes

The sanction is the least discussed term in Becker's framework, but it is where enforcement most often fails. The EU framework deliberately combines two routes. MAR requires member states to give regulators administrative sanctions, with maximum fines of at least three times the profit gained or loss avoided and, for insider dealing and market manipulation, of at least EUR 5 million for individuals and EUR 15 million or 15% of annual turnover for legal persons (Regulation (EU) No 596/2014, 2014), while Directive 2014/57/EU requires criminal sanctions for serious cases (Directive 2014/57/EU, 2014). The European evidence that custodial sentences predict detection (Cumming, Groh, et al., 2018) supports keeping the criminal route. But where criminal cases cannot be proved, as Serbian experience suggests (Kostić & Jelisavac Trošić, 2020), an administrative route with a lower standard of proof allows the regulator to sanction the many cases that fall short of the criminal threshold, reserving prosecution for the most serious, and confiscation of proceeds removes the gain itself.

6.4. Surveillance, exchange rules, and market design

Surveillance is the step between rules and enforcement. Cumming and Johan (2008) survey surveillance in 25 jurisdictions and find that exchanges acting as self-regulatory organisations conduct a wider range of single-market surveillance than securities commissions, and that information-sharing arrangements strengthen surveillance across markets; at the Australian Stock Exchange, sustained surveillance effort was followed by improved integrity outcomes (Comerton-Forde & Rydge, 2006). Aitken, Cumming, and Zhan (2015a) study 22 exchanges from

2003 to 2011 and find that more detailed trading rules and surveillance reduce the number of suspected insider trading cases but increase the profit per suspected case. Deterrence works on the many small opportunistic cases; those that remain are larger and more deliberate, and they call for investigative depth rather than more alerts.

Market design can lower the gains from manipulation directly. Hillion and Suominen (2004) show that a closing call auction reduces closing price manipulation, and Pagano and Schwartz (2003) show that the introduction of a closing call improved market quality on Euronext Paris. Liquidity is itself a defence: Aitken, Cumming, and Zhan (2015b) find that high-frequency trading has significantly reduced the frequency and severity of end-of-day price dislocation, by more than trading rules, surveillance, and enforcement. Measurement matters too: errors in the dates on which colocation was introduced, averaging just over a year, change the measured effects of algorithmic trading on market quality (Aitken, Cumming, & Zhan, 2023; see also Carè & Cumming, 2024). Exchange rules also shape where trading takes place: the share of a cross-listed stock's trading on an exchange rises, at a decreasing rate, with the strength of that exchange's trading rules (Cumming, Hou, & Wu, 2018).

7. Serbia and the region: what transfers and what does not

7.1. The regional setting

Serbia's capital market is small and its supervisor is thinly staffed. Equity market capitalisation was less than 8.3% of gross domestic product in 2022, and the Securities Commission had 38 staff, which the World Bank judged insufficient for the country's capital market strategy (World Bank, 2023). The 2025 mutual evaluation by the Council of Europe's anti-money laundering committee (MONEYVAL) found that staffing constraints have affected the Securities Commission's work since 2022 (MONEYVAL, 2026). The European Commission (2025) assesses Serbia as moderately prepared in financial services and asks it to improve its track record in high-level corruption cases, including the seizure and confiscation of criminal assets and the systematic use of parallel financial investigations. Regional evidence points to an enforcement gap. Kostić and Jelisavac Trošić (2020) find that serious market abuse is criminalised in Serbia but has in practice appeared impossible to prove in court: at the time of their study, special investigative techniques were not available for these offences, and criminal complaints filed by the Securities Commission had not been pursued by prosecutors. Accrual-based earnings management is widespread among 215 large Serbian companies (Miletić & Vučković Milutinović, 2020), and a survey of auditors in Serbia, Croatia, North Macedonia, and Bosnia and Herzegovina finds that Serbian auditors encounter unfounded revenue overstatement more often than auditors in the other countries and report the highest scores for inventory manipulation, and that wider legal powers for auditors would reduce the limits on fraud detection (Dimitrijević et al., 2021).

The trading environment is also changing. Bonds account for about 94% of exchange turnover (Belgrade Stock Exchange, 2026), so surveillance of bond trading deserves as much attention as share manipulation. In 2005–2009 liquidity on the Belgrade market was low and persistent, and swings in foreign investor participation drove sharp changes in illiquidity (Živković & Minović, 2010). In November 2025 the exchange began trading on the OASIS platform of the Athens Exchange Group (Belgrade Stock Exchange, 2025). The rules are moving as well: the EU Listing Act amended MAR, with new provisions on the disclosure of inside information in protracted processes applying from June 2026 (Regulation (EU) 2024/2809, 2024), so Serbia is aligning with a moving target. Neighbouring countries face similar choices: North Macedonia adopted new laws on financial instruments and on prospectuses and transparency in 2024, which establish a Capital Market Supervisory Authority as successor to its Securities Commission (Jovanoska, 2025). The region has also seen what a large reporting failure costs. When the Croatian conglomerate Agrokor ran into crisis, a special law on extraordinary administration of systemically important companies came into force in April 2017 (Law on Extraordinary Administration, 2017), and the re-audited accounts showed HRK 3.9 billion of previously undisclosed liabilities (Agrokor, 2017). Event-study evidence shows that the crisis hit the returns of Agrokor's listed companies but not, significantly, those of other Zagreb stocks (Škrinjaric & Orlović, 2019).

Table 4: From evidence to policy in Serbia: regulatory design levers, the supporting evidence from mature markets, recommended first steps, and the bodies that would lead them

Design lever	Evidence from mature markets	Recommended first step in Serbia	Suggested lead bodies
Credible enforcement	First prosecutions, not new laws, lower the cost of equity.	Select a few cases by deterrent value and strength of evidence; publish outcomes.	Securities Commission; public prosecutors
Sanctions	MAR pairs administrative fines with criminal sanctions; custodial sentences predict detection.	Use the administrative route where the criminal standard cannot be met; confiscate proceeds.	Securities Commission; public prosecutors; courts
Procedural tools	Market abuse has proved hard to prove without investigative powers.	Make special investigative measures and parallel financial investigations available.	Legislator; public prosecutors; police
Supervisory capacity	Regulators' staff and budgets matter; strong regulators outperform weak courts.	Staff the Securities Commission to match its mandate.	Government; Securities Commission

Design lever	Evidence from mature markets	Recommended first step in Serbia	Suggested lead bodies
Detection channels	Employees, media, and whistleblowers uncover fraud; suspicious transaction reports are required under MAR.	Use suspicious transaction and order reports; extend whistleblower channels to market abuse.	Securities Commission; investment firms; Belgrade Stock Exchange
Cooperation	Formal cooperation between authorities predicts detection.	Agree information-sharing and referral protocols.	Securities Commission; prosecutors; police; Tax Administration; Administration for the Prevention of Money Laundering
Surveillance and market design	Algorithmic surveillance, closing auctions, and liquidity reduce manipulation.	Calibrate alerts to few liquid shares and bonds; monitor the closing auction.	Belgrade Stock Exchange; Securities Commission
Data	Even widely used US databases contain errors and omissions.	Publish case-level enforcement data for research.	Securities Commission; courts; universities

Note: Each row links a regulatory lever to the evidence reviewed in Sections 3 to 6 and summarised in Tables 1 and 2. First steps and lead bodies are the author's recommendations, not descriptions of current practice. MAR = EU Market Abuse Regulation (Regulation (EU) No 596/2014); suspicious transaction and order reports are reports that brokers and trading venues must file with the regulator when they suspect market abuse.

Source: Author's synthesis of the studies in Tables 1 and 2 and Sections 6 and 7

7.2. Lessons that transfer

Three lessons travel. The first is that credibility comes before detail. The evidence on first prosecutions (Bhattacharya & Daouk, 2002, 2009) implies that a small number of well-prepared and visible cases, with outcomes made public, does more for market confidence than an elaborate rulebook. Cases should be chosen for their deterrent value and the strength of the evidence rather than their size alone, and prosecutors and the Securities Commission should agree in advance which referrals will be pursued. The second lesson is that detection depends on many actors and on cooperation between them. Whistleblower channels that explicitly cover capital market misconduct, suspicious transaction and order reports from brokers and the exchange (required in the EU under Article 16 of MAR), audit quality, and forensic accounting capacity are the domestic equivalents of the village described by Dyck et al. (2010), and the parallel financial investigations requested by the European

Commission (2025) are where forensic accountants add most value. Formal cooperation predicts detection in Europe (Cumming, Groh, et al., 2018), and information sharing strengthens surveillance (Cumming & Johan, 2008). Where the police, the Tax Administration, the prosecution service, the Administration for the Prevention of Money Laundering, and the Securities Commission each hold part of the evidence, structured cooperation is likely to be the cheapest improvement available. The third lesson is that enforcement capacity matters as much as rules. Supervisory staff predict detection (Cumming, Groh, et al., 2018), public enforcement resources matter (Jackson & Roe, 2009), and a well-resourced regulator outperformed reliance on the courts in Poland (Glaeser et al., 2001). For Serbia this means staffing the Securities Commission to match its mandate, closing the procedural gap identified by Kostić and Jelisavac Trošić (2020) so that investigators have the tools needed to prove market abuse, and using the administrative route where the criminal route cannot succeed.

7.3. Lessons that need adaptation

Other findings need adaptation. In thin markets manipulation is cheap, because a few trades move the price, and alerts calibrated on liquid markets either misfire or miss the pattern. Whatever surveillance software is used, the practical questions are how to calibrate its alerts to a market with few liquid shares and who has the staff to follow them up. Surveillance in a small market can instead track every account in the few actively traded shares, with particular attention to closing prices and to brokers trading for their own account, which is where Khwaja and Mian (2005) found the rents. Closing auctions help but are not a complete answer: on the Athens exchange, whose trading platform Belgrade now uses, a closing call auction reduced but did not eliminate closing price manipulation, and manipulators adapted their strategies (Alexakis, Pappas, & Skarmas, 2021). Designated market makers narrowed bid–ask spreads for most stocks that introduced them on the Zagreb and Ljubljana exchanges, although effects on turnover were inconclusive (Matek & Galić, 2024). Insider trading also appears to be more profitable where enforcement is weak: insider purchases earn abnormal returns on the Zagreb Stock Exchange (Kuveždić & Dedi, 2022), and insider profits in Warsaw were substantially higher than in mature markets, which the authors attribute to insufficient enforcement (Wisniewski & Bohl, 2005).

Concentrated ownership shifts the main risk from fraud against dispersed shareholders towards the expropriation of minority shareholders through related-party transactions, or tunnelling (Johnson, La Porta, Lopez-de-Silanes, & Shleifer, 2000), and towards political connections (Faccio, 2006). Recent international evidence shows that politically connected firms are more likely to win government contracts and contracts of greater value (Cumming, Javakhadze, & Suleymanov, 2024), which links capital market enforcement to the procurement and corruption cases that prosecutors already handle. Private enforcement is weak where class actions are rare and courts are slow, which raises the value of public enforcement

and of the regulator's resources (Cumming, Knill, et al., 2015; Jackson & Roe, 2009). Short sellers and analysts are largely absent, so their role must be played by auditors, employees, and the media. Finally, alignment with MAR is necessary but not sufficient: identical rules produce different outcomes depending on supervisory staff, sanctions, and cooperation (Christensen et al., 2016; Cumming, Groh, et al., 2018).

7.4. A research agenda for the region

The largest gap is empirical. Detected misconduct is not committed misconduct, and even the US databases most widely used in research miss relevant events, record late revelation dates, and include events that are not fraud (Karpoff et al., 2017). Regional research should therefore build hand-verified data from the start, and much of the raw material for large samples already exists. Table 5 lists the main sources, what each contains, and how a researcher can obtain it.

Table 5: Data sources for large-sample research on fraud, manipulation, and enforcement in Serbia and the region: what each contains, how to obtain it, and research uses

Source (website)	Contents	Access	Research use
<i>Serbia</i>			
Financial Statements Register, Serbian Business Registers Agency (apr.gov.rs)	Annual financial statements of companies, banks, insurers, and other entities, from 1994 for most entities	Public on the Agency's website as individual documents; data services on request	Accrual measures, M-score, F-score, and Benford tests on the full population
Business register and register of beneficial owners, Serbian Business Registers Agency	Company status, owners, directors; beneficial owners	Searchable free online after eID registration; extracts for a fee	Related parties, tunnelling, political connections
Belgrade Stock Exchange (bgdx.rs)	Daily to annual trading statistics, indices, block and reported trades, market capitalisation; issuer announcements	Free download with attribution; other data under the exchange's tariff	Liquidity, closing prices, event studies of insider trades and announcements
Central Securities Depository and Clearing House (crhov.rs)	Shareholder records of listed securities	By agreement	Ownership concentration, insider holdings
Securities Commission (sec.gov.rs)	Annual reports, decisions, and supervisory measures	Published documents; case-level data by agreement	Enforcement record and its market effects

Source (website)	Contents	Access	Research use
Courts' portal (portal.sud.rs)	Case tracking for Serbian courts	Online, case by case	Outcomes and duration of criminal and misdemeanour cases
Public Procurement Portal (jnportal.ujn.gov.rs) and open data portal (data.gov.rs)	Procurement notices, contracts, and bidders; other public datasets	Open, machine-readable	Political connections and procurement fraud
World Bank Enterprise Surveys (microdata.worldbank.org)	Firm-level survey of Serbian firms, including corruption questions (2019 round)	Free after registration and confidentiality agreement	Firm-level experience of corruption
<i>Region and comparators</i>			
Annual Financial Statements Registry, FINA, Croatia (fina.hr)	Financial statement data of Croatian entities from 2002	Free public posting	Cross-country accounting screens
JOLP, AJPES, Slovenia (ajpes.si)	Annual reports of Slovenian entities for the last five years	Free	Cross-country accounting screens
European Securities and Markets Authority (esma.europa.eu)	Annual reports on administrative and criminal sanctions under MAR by member state	Free	Enforcement intensity in EU comparators
Orbis, Bureau van Dijk	Harmonised financial and ownership data across countries	Commercial subscription, often through universities	Cross-country panels

Note: Access conditions are as described on each provider's website or in the cited assessments as of October 2026 and may change; non-public records, such as case-level supervisory data and shareholder records, are best obtained through a formal research agreement. Records can be linked across sources using the company registration number. MAR = EU Market Abuse Regulation; FINA = Croatian Financial Agency; AJPES = Agency of the Republic of Slovenia for Public Legal Records and Related Services; JOLP = public posting of annual reports.

Source: Author's compilation from the providers' websites and Đorđević (n.d.)

Four sources allow full-population studies. The Financial Statements Register of the Serbian Business Registers Agency holds the annual financial statements that companies, banks, insurers, and other entities must file, from 1994 for most entities, and the statements are publicly disclosed on the Agency's website (Serbian Business Registers Agency, n.d.). Because the statements are posted as individual documents

rather than as machine-readable data (see Đorđević, n.d., on the company register), a large sample requires either systematic collection from the website or the Agency's data services, and it is worth agreeing access for research with the Agency at the outset. The Public Procurement Portal publishes procurement notices and contracts in machine-readable form and was rated the most open of the datasets assessed in a recent review of public integrity data in Serbia (Đorđević, n.d.). The Belgrade Stock Exchange publishes daily, weekly, monthly, and annual trading statistics, index data, block trades, and reported off-exchange trades that can be downloaded free of charge with attribution, while other historical data are supplied under the exchange's tariff (Belgrade Stock Exchange, n.d.). For neighbouring countries, Croatia's Annual Financial Statements Registry offers free access to financial statement data from 2002 (FINA, n.d.), and Slovenia's AJPES offers free access to the annual reports of the last five years, including more than 75,000 company reports for 2025 (AJPES, n.d.). Commercial databases such as Orbis (Bureau van Dijk) harmonise financial and ownership data across countries and are usually available through university subscriptions.

Enforcement and ownership data require more work but are the most valuable. Securities Commission decisions and annual reports, case records in the courts' portal, and shareholder records held by the Central Securities Depository and Clearing House can be linked by company registration number to build the enforcement panel described below; access to non-public records is best agreed through a formal research cooperation. For comparisons with European Union members, the European Securities and Markets Authority publishes annual reports on the administrative and criminal sanctions imposed under MAR in each member state (European Securities and Markets Authority [ESMA], 2025). Survey data complement these records: the World Bank Enterprise Surveys, conducted in Serbia with the European Bank for Reconstruction and Development and the European Investment Bank, include questions on corruption and are available at firm level free of charge after registration (World Bank, 2019).

With these sources, six projects stand out. First, a panel of Securities Commission decisions, criminal complaints, and court outcomes would measure the enforcement record, the use of administrative and criminal routes, the time to resolution, and the market reaction to enforcement, with comparisons to Croatia, Slovenia, and other EU members using the ESMA data. Second, insider transaction notifications and issuer announcements on the Belgrade Stock Exchange, combined with daily prices, would allow a replication of Kuveždić and Dedi (2022) before and after the 2021 law started to apply in 2023. Third, trade data from Belgrade and its neighbours would allow closing price manipulation measures (Comerton-Forde & Putniņš, 2011) and tests of the effect of the new trading platform and closing mechanisms. Fourth, the full population of financial statements in the Financial Statements Register, with the Croatian and Slovenian registries as comparators, would allow tests of whether accrual-based measures, M-scores, F-scores, and machine-learning screens predict qualified audit opinions, bankruptcy, or enforcement, extending Miletić and Vučković Milutinović (2020) from 215 large companies to the whole

economy. Fifth, procurement records from the Public Procurement Portal matched by registration number to company accounts and ownership would allow tests of political connections and procurement fraud. Sixth, hand-coding the trading rules of Western Balkan exchanges into the indices of Cumming et al. (2011) would show whether staggered alignment with EU rules has improved liquidity, using the natural experiments that accession provides; such work should record both when rules were adopted and when they were implemented with surveillance, since the choice of dates matters for how reforms are evaluated (Cumming & Johan, 2019). For practical advice on publishing such work, see Cumming (2021); when choosing outlets, it is also worth knowing that within one large publisher's coordinated finance-journal ecosystem, citation gains were largest in four journals (Alexander & Cumming, 2026), which also account for about 95% of the estimated revenue gain, mostly through expanded publication volume (Cumming, 2026).

Conclusion

The evidence from mature markets offers Serbia and its neighbours three robust lessons. Enforcement, not the text of the law, is what changes behaviour and the cost of capital. Detection relies on a wide set of actors and on cooperation between them. Enforcement capacity, meaning staff, procedural tools, and workable sanctions, matters as much as the rules themselves. Surveillance and market design can lower the gains from manipulation, but only when calibrated to local conditions: thin trading, concentrated ownership, and weak private enforcement change where misconduct concentrates and which tools work, and rules adopted during EU accession will deliver results only to the degree that they are enforced. The regional evidence, from difficulties in proving market abuse in Serbian courts to the Agrokor case, points in the same direction.

For practitioners, the implication is to give the Securities Commission and prosecutors the staff and procedural tools that market abuse cases require, to concentrate scarce resources on credible cases, to invest in the channels that bring information to investigators, and to treat cooperation across agencies as part of enforcement rather than an addition to it. For researchers, the region is a largely unexplored laboratory. Careful datasets on enforcement, disclosure, trading, and detected fraud in the Western Balkans would fill a real gap in a literature written almost entirely from the vantage point of the largest markets.

Declaration on the use of generative AI

In preparing this paper, the author used Claude, a generative artificial intelligence assistant developed by Anthropic, to search for and summarise relevant literature, edit text, check bibliographic details against publishers' and regulators' records, and assist with summary tables and figures from the published data cited in their sources. The author set the scope and argument of the paper, reviewed and revised all text, tables, and figures produced with the tool, and takes full responsibility for the content of the publication. The tool is not an author of this paper.

References

- Aggarwal, R. K., & Wu, G. (2006). Stock market manipulations. *The Journal of Business*, 79(4), 1915–1953. <https://doi.org/10.1086/503652>
- Agrokor. (2017). *Monthly report: September 2017*. Retrieved on October 6, 2026, from: <https://nagodba.agrokor.hr/storage/2017/12/Agrokor-monthly-report-September-2017-1.pdf>
- Aitken, M., Cumming, D., & Zhan, F. (2015a). Exchange trading rules, surveillance and suspected insider trading. *Journal of Corporate Finance*, 34, 311–330. <https://doi.org/10.1016/j.jcorpfin.2015.07.013>
- Aitken, M., Cumming, D., & Zhan, F. (2015b). High frequency trading and end-of-day price dislocation. *Journal of Banking & Finance*, 59, 330–349. <https://doi.org/10.1016/j.jbankfin.2015.06.011>
- Aitken, M., Cumming, D., & Zhan, F. (2023). Algorithmic trading and market quality: International evidence of the impact of errors in colocation dates. *Journal of Banking & Finance*, 151, 106843. <https://doi.org/10.1016/j.jbankfin.2023.106843>
- AJPES. (n.d.). *JOLP: Public posting of annual reports*. Agency of the Republic of Slovenia for Public Legal Records and Related Services. Retrieved on October 6, 2026, from: <https://ajpes.si/jolp/default.asp>
- Akter, M., Cumming, D., & Ji, S. (2023). Natural disasters and market manipulation. *Journal of Banking & Finance*, 153, 106883. <https://doi.org/10.1016/j.jbankfin.2023.106883>
- Alexakis, C., Pappas, V., & Skarmas, E. (2021). Market abuse under different close price determination mechanisms: A European case. *International Review of Financial Analysis*, 74, 101707. <https://doi.org/10.1016/j.irfa.2021.101707>
- Alexander, C., & Cumming, D. (Eds.). (2020). *Corruption and fraud in financial markets: Malpractice, misconduct and manipulation*. John Wiley & Sons.
- Alexander, C., & Cumming, D. (2026). Coordinated journals, concentrated networks and citation growth: Evidence from finance. *Journal of Banking & Finance*, 188, 107718. <https://doi.org/10.1016/j.jbankfin.2026.107718>
- Allen, F., & Gale, D. (1992). Stock-price manipulation. *The Review of Financial Studies*, 5(3), 503–529. <https://doi.org/10.1093/rfs/5.3.503>
- Amiram, D., Bozanic, Z., Cox, J. D., Dupont, Q., Karpoff, J. M., & Sloan, R. (2018). Financial reporting fraud and other forms of misconduct: A multidisciplinary review of the literature. *Review of Accounting Studies*, 23(2), 732–783. <https://doi.org/10.1007/s11142-017-9435-x>

- Association of Certified Fraud Examiners. (2024). *Occupational fraud 2024: A report to the nations*. Retrieved on October 6, 2026, from: <https://www.acfe.com/-/media/files/acfe/pdfs/rtn/2024/2024-report-to-the-nations.pdf>
- Bao, Y., Ke, B., Li, B., Yu, Y. J., & Zhang, J. (2020). Detecting accounting fraud in publicly traded U.S. firms using a machine learning approach. *Journal of Accounting Research*, 58(1), 199–235. <https://doi.org/10.1111/1475-679X.12292>
- Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2), 169–217. <https://doi.org/10.1086/259394>
- Belgrade Stock Exchange. (2025, November 24). [News release on the first trading day on the OASIS trading platform of the Athens Exchange Group]. Retrieved on October 6, 2026, from: https://www.bgdx.rs/eng/o_berzi/vesti/103021
- Belgrade Stock Exchange. (2026). *Annual report for 2025*. Retrieved on October 6, 2026, from: <https://bgdx.rs/eng/trgovanje/pregled/godisnji/12651>
- Belgrade Stock Exchange. (n.d.). *Historical data*. Retrieved on October 6, 2026, from: https://www.bgdx.rs/eng/trgovanje/istorijski_podaci
- Beneish, M. D. (1999). The detection of earnings manipulation. *Financial Analysts Journal*, 55(5), 24–36. <https://doi.org/10.2469/faj.v55.n5.2296>
- Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society*, 78(4), 551–572.
- Bhattacharya, U., & Daouk, H. (2002). The world price of insider trading. *The Journal of Finance*, 57(1), 75–108. <https://doi.org/10.1111/1540-6261.00416>
- Bhattacharya, U., & Daouk, H. (2009). When no law is better than a good law. *Review of Finance*, 13(4), 577–627. <https://doi.org/10.1093/rof/rfp011>
- Call, A. C., Martin, G. S., Sharp, N. Y., & Wilde, J. H. (2018). Whistleblowers and outcomes of financial misrepresentation enforcement actions. *Journal of Accounting Research*, 56(1), 123–171. <https://doi.org/10.1111/1475-679X.12177>
- Carè, R., & Cumming, D. (2024). Technology and automation in financial trading: A bibliometric review. *Research in International Business and Finance*, 71, 102471. <https://doi.org/10.1016/j.ribaf.2024.102471>
- Chen, J., Cumming, D., Hou, W., & Lee, E. (2013). Executive integrity, audit opinion, and fraud in Chinese listed firms. *Emerging Markets Review*, 15, 72–91. <https://doi.org/10.1016/j.ememar.2012.12.003>
- Chen, J., Cumming, D., Hou, W., & Lee, E. (2016). Does the external monitoring effect of financial analysts deter corporate fraud in China? *Journal of Business Ethics*, 134(4), 727–742. <https://doi.org/10.1007/s10551-014-2393-3>

- Christensen, H. B., Hail, L., & Leuz, C. (2016). Capital-market effects of securities regulation: Prior conditions, implementation, and enforcement. *The Review of Financial Studies*, 29(11), 2885–2924. <https://doi.org/10.1093/rfs/hhw055>
- Coffee, J. C., Jr. (2007). Law and the market: The impact of enforcement. *University of Pennsylvania Law Review*, 156(2), 229–311.
- Comerton-Forde, C., & Putniņš, T. J. (2011). Measuring closing price manipulation. *Journal of Financial Intermediation*, 20(2), 135–158. <https://doi.org/10.1016/j.jfi.2010.03.003>
- Comerton-Forde, C., & Putniņš, T. J. (2014). Stock price manipulation: Prevalence and determinants. *Review of Finance*, 18(1), 23–66. <https://doi.org/10.1093/rof/rfs040>
- Comerton-Forde, C., & Rydge, J. (2006). Market integrity and surveillance effort. *Journal of Financial Services Research*, 29(2), 149–172. <https://doi.org/10.1007/s10693-006-5923-z>
- Commission Delegated Regulation (EU) 2016/522 of 17 December 2015 supplementing Regulation (EU) No 596/2014 of the European Parliament and of the Council. (2016). *Official Journal of the European Union*, L 88, 1–18.
- Commodity Futures Trading Commission. (2008, July 24). [Press release 5521-08 on charges against Optiver and three of its employees]. Retrieved on October 6, 2026, from: <https://www.cftc.gov/PressRoom/PressReleases/pr5521-08>
- Commodity Futures Trading Commission. (2012, April 19). [Press release 6239-12 on the settlement with Optiver US LLC and its traders]. Retrieved on October 6, 2026, from: <https://www.cftc.gov/PressRoom/PressReleases/6239-12>
- Commodity Futures Trading Commission. (2015, April 21). [Press release 7156-15 on charges against Navinder Singh Sarao]. Retrieved on October 6, 2026, from: <https://www.cftc.gov/PressRoom/PressReleases/pr7156-15>
- Cong, L. W., Li, X., Tang, K., & Yang, Y. (2023). Crypto wash trading. *Management Science*, 69(11), 6427–6454. <https://doi.org/10.1287/mnsc.2021.02709>
- Congressional Research Service. (2002). [Report No. RS21253 on the WorldCom accounting restatement]. Retrieved on October 6, 2026, from: <https://www.everycrsreport.com/reports/RS21253.html>
- Corbet, S., Cumming, D., Lucey, B., Peat, M., & Vigne, S. (2020). The destabilising effects of cryptocurrency cybercriminality. *Economics Letters*, 191, 108741. <https://doi.org/10.1016/j.econlet.2019.108741>
- Cressey, D. R. (1953). *Other people's money: A study in the social psychology of embezzlement*. Free Press.

Cumming, D. (2021). What does it take? Tips on research and publishing at the 25th anniversary of the Journal of Corporate Finance. *Journal of Corporate Finance*, 66, 101861. <https://doi.org/10.1016/j.jcorpfin.2020.101861>

Cumming, D. (2026). *The revenue of finance journals: Networks, pricing power, and publication volume* [Working paper]. arXiv. <https://arxiv.org/abs/2508.14301>

Cumming, D., Dannhauser, R., & Johan, S. (2015). Financial market misconduct and agency conflicts: A synthesis and future directions. *Journal of Corporate Finance*, 34, 150–168. <https://doi.org/10.1016/j.jcorpfin.2015.07.016>

Cumming, D., Dannhauser, R., & Johan, S. (2021). Reputational effects of noncompliance with financial market regulations. In B. van Rooij & D. D. Sokol (Eds.), *The Cambridge handbook of compliance* (pp. 245–276). Cambridge University Press. <https://doi.org/10.1017/9781108759458.019>

Cumming, D., Firth, C., Gathergood, J., & Stewart, N. (2023). Work-from-home and the risk of securities misconduct. *European Financial Management*, 29(4), 1054–1077. <https://doi.org/10.1111/eufm.12426>

Cumming, D., Groh, A. P., & Johan, S. (2018). Same rules, different enforcement: Market abuse in Europe. *Journal of International Financial Markets, Institutions and Money*, 54, 130–151. <https://doi.org/10.1016/j.intfin.2018.03.006>

Cumming, D., Hornuf, L., Karami, M., & Schweizer, D. (2023). Disentangling crowdfunding from fraudfunding. *Journal of Business Ethics*, 182(4), 1103–1128. <https://doi.org/10.1007/s10551-021-04942-w>

Cumming, D., Hou, W., & Wu, E. (2018). Exchange trading rules, governance, and trading location of cross-listed stocks. *The European Journal of Finance*, 24(16), 1453–1484. <https://doi.org/10.1080/1351847X.2015.1089522>

Cumming, D., Javakhadze, D., & Suleymanov, M. (2024). *Political connections and government contracting: An international analysis of procurement decisions and firm value* (ECGI Finance Working Paper No. 976/2024). European Corporate Governance Institute. <https://www.ecgi.global/publications/working-papers/political-connections-and-government-contracting-an-international>

Cumming, D., Ji, S., Johan, S., & Tarsalewska, M. (2020). End-of-day price manipulation and M&As. *British Journal of Management*, 31(1), 184–205. <https://doi.org/10.1111/1467-8551.12374>

Cumming, D., Ji, S., Peter, R., & Tarsalewska, M. (2020). Market manipulation and innovation. *Journal of Banking & Finance*, 120, 105957. <https://doi.org/10.1016/j.jbankfin.2020.105957>

Cumming, D., & Johan, S. (2008). Global market surveillance. *American Law and Economics Review*, 10(2), 454–506. <https://doi.org/10.1093/aler/ahn009>

- Cumming, D., & Johan, S. (2013). Listing standards and fraud. *Managerial and Decision Economics*, 34(7–8), 451–470. <https://doi.org/10.1002/mde.2622>
- Cumming, D., & Johan, S. (2019). Capital-market effects of securities regulation: Prior conditions, implementation, and enforcement revisited. *Finance Research Letters*, 31, 425–435. <https://doi.org/10.1016/j.frl.2018.12.013>
- Cumming, D., Johan, S., & Li, D. (2011). Exchange trading rules and stock market liquidity. *Journal of Financial Economics*, 99(3), 651–671. <https://doi.org/10.1016/j.jfineco.2010.10.001>
- Cumming, D., Knill, A., & Richardson, N. (2015). Firm size and the impact of securities regulation. *Journal of Comparative Economics*, 43(2), 417–442. <https://doi.org/10.1016/j.jce.2014.11.003>
- Cumming, D., Leung, T. Y., & Rui, O. (2015). Gender diversity and securities fraud. *Academy of Management Journal*, 58(5), 1572–1593. <https://doi.org/10.5465/amj.2013.0750>
- Danas. (2021, February 9). *Za malverzacije u Agrobanci deset okrivljenih osuđeno na ukupno 23 godine zatvora* [Ten defendants sentenced to a total of 23 years in prison for malfeasance at Agrobanka]. Retrieved on October 6, 2026, from: <https://www.danas.rs/drustvo/za-malverzacije-u-agrobanci-deset-okrivljenih-osudjeno-na-ukupno-23-godine-zatvora/>
- Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G. (2011). Predicting material accounting misstatements. *Contemporary Accounting Research*, 28(1), 17–82. <https://doi.org/10.1111/j.1911-3846.2010.01041.x>
- Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1995). Detecting earnings management. *The Accounting Review*, 70(2), 193–225.
- Del Guercio, D., Odders-White, E. R., & Ready, M. J. (2017). The deterrent effect of the Securities and Exchange Commission's enforcement intensity on illegal insider trading: Evidence from run-up before news events. *Journal of Law and Economics*, 60(2), 269–307. <https://doi.org/10.1086/693563>
- Dey, A., Heese, J., & Pérez-Cavazos, G. (2021). Cash-for-information whistleblower programs: Effects on whistleblowing and consequences for whistleblowers. *Journal of Accounting Research*, 59(5), 1689–1740. <https://doi.org/10.1111/1475-679X.12370>
- Dhawan, A., & Putniņš, T. J. (2023). A new wolf in town? Pump-and-dump manipulation in cryptocurrency markets. *Review of Finance*, 27(3), 935–975. <https://doi.org/10.1093/rof/rfac051>
- Dimitrijević, D., Jovković, B., & Milutinović, S. (2021). The scope and limitations of external audit in detecting frauds in company's operations. *Journal of Financial Crime*, 28(3), 632–646. <https://doi.org/10.1108/JFC-11-2019-0155>

Dimitrijević, D., Obradović, V., & Milutinović, S. (2018). Indicators of fraud in financial reporting in the Republic of Serbia. *Teme*, 42(4), 1319–1338. <https://doi.org/10.22190/TEME1804319D>

Directive 2014/57/EU of the European Parliament and of the Council of 16 April 2014 on criminal sanctions for market abuse (market abuse directive). (2014). *Official Journal of the European Union*, L 173, 179–189.

Dyck, A., Morse, A., & Zingales, L. (2010). Who blows the whistle on corporate fraud? *The Journal of Finance*, 65(6), 2213–2253. <https://doi.org/10.1111/j.1540-6261.2010.01614.x>

Dyck, A., Morse, A., & Zingales, L. (2024). How pervasive is corporate fraud? *Review of Accounting Studies*, 29(1), 736–769. <https://doi.org/10.1007/s11142-022-09738-5>

Enron Corp. (2001, November 8). *Form 8-K: Current report*. U.S. Securities and Exchange Commission. Retrieved on October 6, 2026, from: <https://www.sec.gov/Archives/edgar/data/0001024401/000095012901503835/h91831e8-k.txt>

European Commission. (2025). *Serbia 2025 report* (Commission Staff Working Document SWD(2025) 755 final). Retrieved on October 6, 2026, from: <https://www.stat.gov.rs/media/410862/serbia-report-2025.pdf>

European Securities and Markets Authority. (2025). [Annual report on sanctions and measures imposed by national competent authorities] (ESMA43-1527801302-1828). Retrieved on October 6, 2026, from: https://www.esma.europa.eu/sites/default/files/2025-10/ESMA43-1527801302-1828_Annual_Sanctions_Report_2025.pdf

Faccio, M. (2006). Politically connected firms. *American Economic Review*, 96(1), 369–386. <https://doi.org/10.1257/000282806776157704>

FINA. (n.d.). *Annual Financial Statements Registry (RGFI)*. Croatian Financial Agency. Retrieved on October 6, 2026, from: <https://www.fina.hr/eng/public-services-business/registries/annual-financial-statements-registry-rgfi>

Financial Conduct Authority. (2021). *Market Watch 67*. Retrieved on October 6, 2026, from: <https://www.fca.org.uk/publications/newsletters/market-watch-67>

Financial Services Authority. (2012). *Final notice: Barclays Bank plc*. Retrieved on October 6, 2026, from: <https://www.fca.org.uk/publication/final-notice/barclays-jun12.pdf>

Giannetti, M., & Wang, T. Y. (2016). Corporate scandals and household stock market participation. *The Journal of Finance*, 71(6), 2591–2636. <https://doi.org/10.1111/jofi.12399>

- Glaeser, E., Johnson, S., & Shleifer, A. (2001). Coase versus the Coasians. *The Quarterly Journal of Economics*, 116(3), 853–899. <https://doi.org/10.1162/00335530152466250>
- Hillion, P., & Suominen, M. (2004). The manipulation of closing prices. *Journal of Financial Markets*, 7(4), 351–375. <https://doi.org/10.1016/j.finmar.2004.04.002>
- Isaković-Kaplan, Š., Demirović, L., & Proho, M. (2021). Benford's law in forensic analysis of income statements of economic entities in Bosnia and Herzegovina. *Croatian Economic Survey*, 23(1), 31–61. <https://doi.org/10.15179/ces.23.1.2>
- Jackson, H. E., & Roe, M. J. (2009). Public and private enforcement of securities laws: Resource-based evidence. *Journal of Financial Economics*, 93(2), 207–238. <https://doi.org/10.1016/j.jfineco.2008.08.006>
- Johnson, S., La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2000). Tunneling. *American Economic Review*, 90(2), 22–27. <https://doi.org/10.1257/aer.90.2.22>
- Jones, J. J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research*, 29(2), 193–228. <https://doi.org/10.2307/2491047>
- Jovanoska, A. (2025, March 14). From 2025 to 2026: Banking and finance in North Macedonia. *CEE Legal Matters*. Retrieved on October 6, 2026, from: <https://ceelm.com/magazine-articles/11053-issue-13-1/32196-from-2025-to-2026-banking-and-finance-in-north-macedonia>
- Karpoff, J. M., Koester, A., Lee, D. S., & Martin, G. S. (2017). Proxies and databases in financial misconduct research. *The Accounting Review*, 92(6), 129–163. <https://doi.org/10.2308/accr-51766>
- Karpoff, J. M., Lee, D. S., & Martin, G. S. (2008a). The consequences to managers for financial misrepresentation. *Journal of Financial Economics*, 88(2), 193–215. <https://doi.org/10.1016/j.jfineco.2007.06.003>
- Karpoff, J. M., Lee, D. S., & Martin, G. S. (2008b). The cost to firms of cooking the books. *Journal of Financial and Quantitative Analysis*, 43(3), 581–611. <https://doi.org/10.1017/S0022109000004221>
- Karpoff, J. M., & Lou, X. (2010). Short sellers and financial misconduct. *The Journal of Finance*, 65(5), 1879–1913. <https://doi.org/10.1111/j.1540-6261.2010.01597.x>
- Khwaja, A. I., & Mian, A. (2005). Unchecked intermediaries: Price manipulation in an emerging stock market. *Journal of Financial Economics*, 78(1), 203–241. <https://doi.org/10.1016/j.jfineco.2004.06.014>
- Kostić, J., & Jelisavac Trošić, S. (2020). (In)adequate criminal protection of the capital market in the Republic of Serbia. *EU and Comparative Law Issues and Challenges Series (ECLIC)*, 4, 598–620. <https://doi.org/10.25234/eclic/11918>

Kuveždić, M., & Dedi, L. (2022). Insider trading at Zagreb Stock Exchange. *Zagreb International Review of Economics and Business*, 25(1), 79–94. <https://doi.org/10.2478/zireb-2022-0006>

La Porta, R., Lopez-de-Silanes, F., & Shleifer, A. (2006). What works in securities laws? *The Journal of Finance*, 61(1), 1–32. <https://doi.org/10.1111/j.1540-6261.2006.00828.x>

Law on Capital Market [Zakon o tržištu kapitala]. (2021). *Službeni glasnik Republike Srbije*, No. 129/2021 and No. 109/2025.

Law on Digital Assets [Zakon o digitalnoj imovini]. (2020). *Službeni glasnik Republike Srbije*, No. 153/2020.

Law on Extraordinary Administration [Zakon o postupku izvanredne uprave u trgovačkim društvima od sistemskog značaja za Republiku Hrvatsku]. (2017). *Narodne novine*, No. 32/2017.

Lennox, C., & Pittman, J. A. (2010). Big Five audits and accounting fraud. *Contemporary Accounting Research*, 27(1), 209–247. <https://doi.org/10.1111/j.1911-3846.2010.01007.x>

Matek, P.-P., & Galić, M. (2024). The impact of designated market-makers on liquidity in frontier markets: Evidence from Zagreb and Ljubljana Stock Exchanges. *Zbornik radova Ekonomskog fakulteta u Rijeci*, 42(1), 95–121. <https://doi.org/10.18045/zbefri.2024.1.95>

Miletić, S., & Vučković Milutinović, S. (2020). An analysis of accrual-based earnings management in large Serbian companies. *Ekonomika preduzeća*, 68(7–8), 469–483. <https://doi.org/10.5937/EKOPRE2008469M>

MONEYVAL. (2026, February 4). *Strides made in risk assessment and prosecution of money laundering in Serbia* [Press release on the 2025 mutual evaluation of Serbia]. Council of Europe. Retrieved on October 6, 2026, from: <https://www.coe.int/en/web/portal/-/strides-made-in-risk-assessment-and-prosecution-of-money-laundering-in-serbia>

Nasdaq. (n.d.). *Market surveillance*. Retrieved on October 6, 2026, from: <https://www.nasdaq.com/products/fintech/market-surveillance>

NASDAQ OMX Group. (2010, July 27). *NASDAQ OMX to acquire world-leading market surveillance system provider SMARTS* [Press release, Exhibit 99.1 to Form 8-K]. Retrieved on October 6, 2026, from: <https://www.sec.gov/Archives/edgar/data/0001120193/000119312510166688/dex991.htm>

National Bank of Serbia. (n.d.). [Web page on bank resolutions in 2012–2014]. Retrieved on October 6, 2026, from: https://nbs.rs/en/o_nbs/jub140/jub_fs/

Nigrini, M. J. (2012). *Benford's law: Applications for forensic accounting, auditing, and fraud detection*. John Wiley & Sons.

Dorđević, M. (n.d.). *Data scoping report for Serbia* [Integrity Watch Western Balkans and Turkey project]. Transparency Serbia. Retrieved on October 6, 2026, from:

https://www.transparentnost.org.rs/images/dokumenti_uz_vesti/Data_Scoping_Report_for_Serbia_-_IW.pdf

Organized Crime and Corruption Reporting Project. (2012). *Serbia: New government cracks down on corrupt bank officials*. Retrieved on October 6, 2026, from: <https://occrp.org/en/news/serbia-new-government-cracks-down-on-corrupt-bank-officials>

Pagano, M. S., & Schwartz, R. A. (2003). A closing call's impact on market quality at Euronext Paris. *Journal of Financial Economics*, 68(3), 439–484. [https://doi.org/10.1016/S0304-405X\(03\)00073-4](https://doi.org/10.1016/S0304-405X(03)00073-4)

Petrović, D., Novović, M., & Šoškić, M. (2025). Comparative analysis of milling-bakery and confectionery industry in Serbia based on Benford's law. *Economics of Agriculture*, 72(1), 13–32. <https://doi.org/10.59267/ekoPolj250113P>

Policy and Legal Advice Centre IV. (2025, October 20). *Workshop on aligning Serbia's regulatory framework with EU rules on preventing capital market abuse*. Retrieved on October 6, 2026, from: <https://euinfo.rs/plac4/?p=829>

Regulation (EU) 2024/2809 of the European Parliament and of the Council of 23 October 2024 amending Regulations (EU) 2017/1129, (EU) No 596/2014 and (EU) No 600/2014 to make public capital markets in the Union more attractive for companies and to facilitate access to capital for small and medium-sized enterprises. (2024). *Official Journal of the European Union*, L series. <https://eur-lex.europa.eu/eli/reg/2024/2809/oj>

Regulation (EU) No 596/2014 of the European Parliament and of the Council of 16 April 2014 on market abuse (market abuse regulation). (2014). *Official Journal of the European Union*, L 173, 1–61.

Serbian Business Registers Agency. (n.d.). *Financial statements register: About the register*. Retrieved on October 6, 2026, from: <https://apr.gov.rs/registers/financial-statements/about-register.1699.html>

Škrinjarić, T., & Orlović, Z. (2019). Effects of economic and political events on stock returns: Event study of the Agrokor case in Croatia. *Croatian Economic Survey*, 21(1), 47–86. <https://doi.org/10.15179/ces.21.1.2>

Spasić, D., & Stojanović, J. (2026). The nexus between fraudulent financial reporting and company failure – The case of bankrupt companies in Serbia. *Teme*, 50(2), 437–460. <https://doi.org/10.22190/TEME260131028S>

U.S. Department of Justice. (2020, September 29). [Press release on JPMorgan Chase & Co. agreeing to pay \$920 million in connection with schemes to defraud precious metals and U.S. Treasuries markets]. Retrieved on October 6, 2026, from: <https://www.justice.gov/opa/pr/jpmorgan-chase-co-agrees-pay-920-million-connection-schemes-defraud-precious-metals-and-us>

U.S. Department of Justice. (2024, October 9). [Press release on eighteen individuals and entities charged in an international operation targeting market manipulation in cryptocurrency]. Retrieved on October 6, 2026, from: <https://www.justice.gov/usao-ma/pr/eighteen-individuals-and-entities-charged-international-operation-targeting-widespread>

U.S. Securities and Exchange Commission. (2003). *Litigation release No. 18527: Parmalat Finanziaria S.p.A.* Retrieved on October 6, 2026, from: <https://www.sec.gov/litigation/litreleases/lr18527.htm>

U.S. Securities and Exchange Commission. (2011a). [Press release 2011-81 on charges against Satyam Computer Services]. Retrieved on October 6, 2026, from: <https://www.sec.gov/news/press/2011/2011-81.htm>

U.S. Securities and Exchange Commission. (2011b). [Press release 2011-82 on charges against India-based affiliates of PricewaterhouseCoopers in connection with Satyam]. Retrieved on October 6, 2026, from: <https://www.sec.gov/news/press/2011/2011-82.htm>

U.S. Securities and Exchange Commission. (2020, December 16). [Press release 2020-319 on settled accounting fraud charges against Luckin Coffee]. Retrieved on October 6, 2026, from: <https://www.sec.gov/news/press-release/2020-319>

United States v. Coscia, 866 F.3d 782 (7th Cir. 2017).

University of New South Wales. (2010, March). *Eye on the market*. UNSW Newsroom. Retrieved on October 6, 2026, from: <https://www.unsw.edu.au/newsroom/news/2010/03/eye-on-the-market>

Wirecard AG. (2020, June 18). [Ad hoc announcement on the delayed publication of the 2019 financial statements due to indications of spurious balance confirmations]. Retrieved on October 6, 2026, from: <https://wirecard.com/2020/06/18/wirecard-ag-date-for-publication-of-annual-and-consolidated-financial-statements-2019-delayed-due-to-indications-of-presentation-of-spurious-balance-confirmations/>

Wisniewski, T. P., & Bohl, M. T. (2005). The information content of registered insider trading under lax law enforcement. *International Review of Law and Economics*, 25(2), 169–185. <https://doi.org/10.1016/j.irle.2005.06.002>

World Bank. (2019). *Serbia Enterprise Survey 2019* [Data set]. World Bank Microdata Library. Retrieved on October 6, 2026, from: <https://microdata.worldbank.org/index.php/catalog/3721>

World Bank. (2023). *Catalyzing Long Term Finance through Capital Markets Project: Project appraisal document* (Report No. PAD4559). Retrieved on October 6, 2026, from: <https://documents1.worldbank.org/curated/en/099212502272311919/pdf/BOSIB09b18d4b20420a8ea0f82a5b662db6.pdf>

Živković, B., & Minović, J. (2010). Illiquidity of frontier financial market: Case of Serbia. *Panoeconomicus*, 57(3), 349–367. <https://doi.org/10.2298/PAN1003349Z>

Saeid Homayoun
University of Gävle
ORCID: 0000-0002-2536-0446

Zabihollah Rezaee
Fogelman College of Business and
Economics, University of Memphis;
Corvinus University of Budapest,
Corvinus Institute for Advanced
Studies
ORCID: 0000-0001-9889-9699

Zoltán Oszkár SZÁNTÓ
Corvinus University of Budapest,
Corvinus Institute for Advanced
Studies
ORCID: 0000-0002-3694-8544

DO ESG-CENTRIC FIRMS ENGAGE IN FRAUDULENT ACTIVITIES?

UDK 657.6:005.35
DOI: <https://doi.org/10.54318/czf.2026.sh.002>

Abstract

This study examines the relationships among environmental and social performance, corporate governance, sustainability reporting, external assurance, and ESG controversy scores. The reported sample contains 12,295 firm-year observations for 2003–2019, including 3,952 sustainability reports and 1,056 assured reports. Environmental and social performance is positively associated with governance performance across the reported specifications. Governance and environmental and social performance are negatively associated with ESG controversy scores, and their interaction is also negative. Reporting and assurance further condition these relationships. Under the provider's conventional scoring direction, higher controversy scores indicate more favorable controversy performance; the negative coefficients therefore do not demonstrate fewer controversies. The evidence highlights the need to distinguish sustainability ratings, disclosure credibility, and observed conduct. The regressions establish associations, not causal effects or findings of fraud. Interpretation remains conditional on confirming the historical data fields, transformations, sample construction, and estimation procedures.

Key words: ESG performance; corporate governance; ESG controversies; sustainability reporting; sustainability assurance; disclosure credibility

JEL classification: M41, M42, G34, Q56

Introduction

Sustainability reporting has expanded the information available to investors and other stakeholders about corporate environmental, social, and governance (ESG) practices. Yet favorable ratings and extensive disclosure do not necessarily establish responsible conduct. Firms may invest in sustainability capabilities while still experiencing environmental violations, labor disputes, product-safety failures, or governance controversies. Reports may communicate substantive achievements, selectively emphasize favorable information, or reflect a combination of these motives. The research question address in our study is “Do firms that issue sustainability report engage less in financial reporting fraud (hereafter, FRF)?” More specifically we examine: (1) the link between sustainability reporting and FRF; (2) the association between both financial economic sustainability performance (ESP) and non-financial environmental, social, and governance (ESG) sustainability performance information and FRF. The central question is whether sustainability performance, governance, reporting, and assurance are associated with the controversies recorded by external information providers.

We examine five related relationships: environmental and social performance with governance performance; governance with controversy scores; environmental and social performance with controversy scores; governance as a moderator of the performance–controversy relationship; and reporting as a moderator of that relationship. We additionally examine external assurance, including a comparison of assured and unassured reports within the reporting subsample. Separating these constructs is essential because governance ratings, environmental and social ratings, disclosure choices, and externally recorded controversies capture different aspects of corporate accountability.

Prior research provides reasons to expect both alignment and divergence. Sustainability activities can become embedded in organizational processes and stakeholder relationships (Eccles et al., 2014; Rezaee, 2016). They may also support risk management and reputational capital (Godfrey et al., 2009; Albuquerque et al., 2019). Conversely, greenwashing and selective disclosure can produce a gap between public claims and organizational conduct (Delmas and Burbano, 2011; Marquis et al., 2016). Neither the existence of a sustainability report nor a high performance score resolves this tension.

The reported sample comprises 12,295 firm-year observations during 2003–2019. Sustainability reports are identified for 3,952 observations, of which 1,056 receive external assurance. The tables identify Asset4 score fields for controversy, governance, environmental, and social performance. Environmental and social performance is positively associated with governance performance in all four governance models. In the controversy models, governance and environmental and social performance generally have negative coefficients, while their interaction and the interactions with reporting and assurance are also negative.

These signs require careful interpretation. The convention documented in the provider's methodology assigns more favorable controversy scores to firms with less recorded controversy exposure. Thus, a negative coefficient on the score cannot be described as evidence of controversy prevention unless the dependent variable was transformed into an increasing measure of controversy exposure. The reported minimum, maximum, and median are consistent with a bounded score, not an incident count. We therefore interpret the estimates at the score level and distinguish the conventional provider interpretation from any transformation that must be verified against the analysis files.

The study contributes to the discussion of sustainability credibility by considering performance, governance, reporting, and assurance together. It also identifies a measurement problem that affects substantive conclusions: the direction of a score determines whether a coefficient supports or contradicts a prevention hypothesis. More broadly, observed controversies may reflect underlying conduct, public visibility, media coverage, and provider classification. An association between high performance ratings and unfavorable controversy scores need not establish managerial deception, and an association with favorable scores need not prove that misconduct was prevented.

The remainder of the paper reviews relevant research and develops the hypotheses, describes the variables and model structure, discusses the reported results, and presents implications and limitations. The analysis is observational. It does not directly test financial reporting fraud, adjudicate the truthfulness of individual sustainability claims, or establish the causal effects of governance, disclosure, or assurance.

1. Literature and theoretical foundations

1.1. Sustainability performance and economic outcomes

Research on sustainability encompasses organizational practices, stakeholder relationships, disclosure, and economic consequences. Accounting reviews organize this literature around the determinants and consequences of CSR, financial performance, reporting, and assurance (Huang and Watson, 2015; Rezaee, 2016). Research synthesis also indicates that estimated social and financial performance relationships depend on the methods and settings examined (Busch and Friede, 2018). These findings motivate examining specific mechanisms rather than treating sustainability as a single, uniformly beneficial attribute.

Materiality is particularly important. Khan et al. (2016) distinguish sustainability investments associated with material industry issues from other investments. This distinction suggests that aggregate scores may obscure differences in the risks most relevant to a firm's operations. A company may perform well on many disclosed indicators while remaining exposed to a material weakness in a particular activity, location, or supply chain. The empirical design must therefore avoid interpreting an

overall performance score as comprehensive protection against all forms of misconduct.

Disclosure also conditions the information available to stakeholders. Fatemi et al. (2018) examine the interaction between ESG performance, disclosure, and valuation, while Servaes and Tamayo (2013) emphasize customer awareness in the CSR–value relationship. Cohen et al. (2015) document professional investors’ preferences for credible and usable nonfinancial information. Together, these studies support separating what a firm does from what stakeholders learn about it. They do not establish that reporting itself prevents misconduct.

1.2. Stakeholders, governance and organizational accountability

Stakeholder theory identifies the interests and relationships that influence corporate decisions and accountability (Donaldson and Preston, 1995). Environmental and social responsibilities concern employees, customers, suppliers, communities, and other affected groups. Governance provides the monitoring and decision processes through which these responsibilities may enter strategy, investment, incentives, and risk management. This paper uses the stakeholder perspective to develop a proposed connection between environmental and social performance and governance capability.

Governance and sustainability may reinforce each other. Boards can allocate responsibility for material risks, review performance information, challenge management, and monitor corrective action. Sustainability initiatives can, in turn, require coordination and oversight that strengthen organizational routines. The organizational-process perspective in Eccles et al. (2014) provides a basis for this argument. However, the existence of a committee, policy, or disclosure system should not be equated with effective implementation. Observable structures and rated performance remain proxies for the underlying quality of oversight.

An agency perspective provides a competing interpretation. Managers may use visible sustainability initiatives to enhance their reputations, secure stakeholder support, or divert attention from weak monitoring. Strong governance could constrain such opportunism, but managers and directors may also disagree about which sustainability expenditures serve the firm’s objectives. The empirical relationship between governance and environmental and social performance is consequently an association to be tested, rather than an assumed identity.

1.3. Legitimacy, signaling and controversy visibility

Legitimacy theory links organizational acceptance to the fit between corporate activities and societal expectations (Dowling and Pfeffer, 1975). Sustainability reports can support legitimacy by explaining commitments and performance. Signaling arguments further suggest that disclosure may convey information about organizational capability or anticipated performance; Lys et al. (2015) examine signaling through corporate accountability reporting. In this study, reporting is

proposed to make CSR commitments more visible and potentially more costly to abandon.

The competing possibility is symbolic communication. Firms can selectively disclose favorable achievements without adequately addressing material weaknesses. Delmas and Burbano (2011) analyze drivers of greenwashing, and Marquis et al. (2016) examine scrutiny, norms, and selective disclosure. These perspectives imply that reporting quantity or a favorable narrative cannot independently establish the reliability of the underlying conduct. Distinguishing substantive accountability from impression management requires evidence about outcomes, reporting boundaries, controls, and contradictory information.

Controversy data introduce a further complication. Public visibility can increase the probability that an adverse event receives attention and enters an ESG database. Barkemeyer et al. (2023) identify selection bias in controversy coverage, while Aouadi and Marsat (2018) show that controversy–valuation relationships depend on the empirical context and interactions examined. Thus, recorded controversies are not a complete census of misconduct. Firms with extensive sustainability communication may face closer scrutiny, making their problems more observable even if their underlying practices are comparatively strong.

Rating disagreement reinforces the need for construct discipline. Berg et al. (2022) document divergence across ESG rating providers, and Christensen et al. (2022) examine disagreement in ESG assessments. A score reflects a provider's coverage, measurement choices, and aggregation procedures. It cannot be treated as interchangeable with another provider's score, a count of adverse events, an adjudicated legal violation, or a measure of intentional fraud.

1.4. Assurance and disclosure credibility

External assurance can enhance confidence in sustainability information by subjecting specified disclosures to independent procedures. Simnett et al. (2009) examine international assurance choices, and Cuadrado-Ballesteros et al. (2017) analyze assurance in relation to information asymmetry. Assurance may also encourage documentation, clearer reporting responsibilities, and review of supporting evidence. These are proposed channels through which assurance could influence the credibility of reporting and its relationship with controversy outcomes. Nevertheless, assurance engagements differ in scope, criteria, provider characteristics, and level of assurance. Critical evaluations identify limitations in assurance statements and managerial influence over the process (O'Dwyer and Owen, 2005; O'Dwyer et al., 2011). Venter and van Eck (2021) synthesize research on extended external reporting assurance. A binary assurance indicator cannot capture these distinctions or demonstrate that all material sustainability activities were independently evaluated.

Reporting and assurance are also selected rather than randomly assigned. Firms may report or obtain assurance because they have better systems, greater resources, more

stakeholder pressure, or greater exposure to adverse events. The review by Christensen et al. (2021) highlights the importance of the economic incentives and implementation conditions surrounding sustainability reporting. Accordingly, associations involving reporting or assurance must be distinguished from causal effects on conduct or disclosure reliability.

2. Hypothesis development

2.1. Environmental and social performance and governance

Environmental and social sustainability requires coordination across operations, human resources, procurement, compliance, and risk management. Firms that develop these capabilities may also strengthen oversight, reporting responsibilities, and accountability. Governance can make sustainability commitments more durable by linking them to resource allocation and managerial incentives. Conversely, substantive sustainability practices can create demand for better governance information and board attention. This proposed complementarity is consistent with an organizational-process interpretation of sustainability.

The prediction must avoid circular measurement. If an overall ESG score includes governance, regressing governance on that score may partly reproduce a shared component. The reported CSR performance measure instead averages environmental and social scores, excluding governance. This separation does not eliminate common-provider measurement dependence, but it reduces direct overlap in the explanatory and dependent constructs. The hypothesis concerns association and does not establish whether sustainability precedes or follows governance improvements.

H1. Environmental and social sustainability performance is positively associated with corporate governance performance.

2.2. Governance and controversies

Effective governance can constrain opportunism and improve the identification, escalation, and remediation of material risks. Board oversight, internal audit, compliance programs, whistleblowing arrangements, and risk-management systems may reduce the conditions that generate adverse environmental or social events. Governance can also support a timely response when an event occurs, potentially limiting its recurrence or severity.

Formal governance structures may, however, be weakly implemented. More transparent governance can also increase detection and disclosure, raising observed controversies without increasing underlying misconduct. The prevention prediction therefore concerns substantive oversight, while the empirical measure remains a governance score. Under the conventional controversy-score direction, prevention implies a positive score coefficient rather than a negative one.

H2. Corporate governance performance is negatively associated with ESG controversy exposure, corresponding to a positive association with a conventionally coded controversy score.

2.3. Environmental and social performance and controversies

Substantive CSR practices may reduce exposure to employee harm, product-safety failures, human-rights abuses, and environmental damage. Relevant activities include workforce training, supplier due diligence, complaint handling, operational controls, and corrective action. Firms with valuable stakeholder relationships may have incentives to avoid conduct that jeopardizes those relationships. Reputational capital and risk-management arguments motivate this proposed prevention mechanism (Barnett, 2007; Godfrey et al., 2009).

An alternative visibility mechanism predicts a weaker or reversed relationship in recorded data. Strong sustainability claims attract scrutiny and create standards against which adverse conduct can be judged. A controversy involving a highly rated firm may receive more attention than a similar event at a less visible firm. Consequently, a negative relationship with the conventionally coded controversy score would challenge the prevention hypothesis but would not, by itself, establish that sustainability engagement causes misconduct.

H3. Environmental and social sustainability performance is negatively associated with ESG controversy exposure, corresponding to a positive association with a conventionally coded controversy score.

2.4. Governance as a moderator

CSR commitments require implementation through operating procedures, incentives, monitoring, and accountability. Governance may supply this infrastructure by assigning responsibility, reviewing material risks, testing controls, and requiring corrective action. When these mechanisms operate effectively, governance should reinforce the preventive association between CSR performance and controversy exposure.

The competing explanation is greater reputational exposure. Firms that combine strong CSR claims with visible governance structures may experience heightened scrutiny or a larger credibility loss after an adverse event. The hypothesis predicts that prevention dominates this mechanism, but the regression interaction can be interpreted only after specifying the dependent variable's direction. Its sign describes how the slope changes; it does not independently identify an operating mechanism.

H4. Governance performance strengthens the negative association between environmental and social performance and controversy exposure, corresponding to a positive interaction in a conventionally coded controversy-score model.

2.5. Sustainability reporting as a moderator

Preparing a sustainability report can require a firm to collect information, assign data responsibilities, evaluate progress, and communicate objectives. Public commitments may encourage internal discipline and permit stakeholders to compare claims with subsequent outcomes. Reporting could therefore strengthen the preventive association of substantive CSR performance with controversy exposure.

Reporting can also increase visibility or serve as selective communication. Reporters may be larger, face greater stakeholder pressure, or have better documentation of adverse events. The empirical reporting indicator captures the existence of a report rather than its balance, materiality, accuracy, or completeness. The prevention hypothesis is therefore distinct from the alternative prediction that reporting intensifies the observed association between favorable CSR ratings and recorded controversy exposure.

H5. Sustainability reporting strengthens the negative association between environmental and social performance and controversy exposure, corresponding to a positive performance–reporting interaction in a conventionally coded controversy-score model.

External assurance is examined as an additional credibility mechanism. We compare its interaction with CSR performance in the full sample and among reporters. The latter comparison is conceptually more direct because the full-sample no-assurance category combines nonreporters with firms that publish unassured reports. We do not treat a binary assurance measure as evidence that every relevant activity or claim was verified.

3. Research method

3.1. Sample and data boundaries

Table 1 reports 12,295 firm-year observations for 2003–2019 across 11 sectors. Sustainability reports are identified for 3,952 observations, or 32.1% of the sample. Of these reports, 1,056 are externally assured, representing 26.7% of reporting observations and 8.6% of the full sample. The sample is concentrated in later years, with 2016–2019 accounting for approximately 53.1% of observations. Changes in coverage and sample composition may therefore influence pooled relationships.

The score fields identified in Table 2 are Asset4 TRESGCS, CGSCORE, ENSCORE, and SOSCORE. The financial variables are described as derived from financial statement data. The identified data sources include Compustat and CRSP. Complete documentation of the geographic universe, firm identifiers, database extraction dates, matching rules, missing-data exclusions, and derivation of the final sample remains necessary for replication. Interpretation is limited to the reported 2003–2019 score-based sample.

3.2. Variable definitions

ESG_Con is the reported controversy score associated with TRESGCS. It is bounded between zero and 100 in the described scoring system. The historical methodology available from LSEG assigns 100 to companies without recorded controversies and penalizes adverse events; the measure is therefore not an increasing incident count (LSEG, 2022). This paper interprets the published field convention while recognizing that the actual extraction and any transformation must be verified. A reverse-coded exposure measure would change coefficient signs without changing the underlying relationship.

CGP is the governance performance score identified as CGSCORE. CSRP is the equally weighted average of ENSCORE and SOSCORE. Despite the CSR label, CSRP combines environmental and social performance; it is not a social-only score or a governance-inclusive overall ESG measure. Using the two pillars separately in future analysis would clarify whether the pooled relationship is concentrated in one dimension.

CSR_Report indicates the presence of a sustainability report, and CSR_Assurance indicates external assurance of a report. We use CSR_Report consistently to denote the reporting indicator. Neither indicator records the scope, framework, provider, assurance level, or quality of the engagement. The reporter-only analysis uses 3,952 observations, while full-sample models use 12,295 observations as printed in the regression tables.

The financial controls are accruals, firm size, cash holdings, leverage, market-to-book value, and sales growth. The table defines Asset as the natural logarithm of total assets, Cash as cash divided by assets, Lev as debt divided by assets, MTB as the market-to-book ratio of assets, and SaleGr as the rate of change in sales or revenue. The accrual measure is described through changes in working-capital components, but its scaling and any depreciation adjustment are not fully documented. It cannot be interpreted as abnormal accruals or a validated fraud measure without additional construction details.

3.3. Model structure

The reported tables contain linear score regressions with lagged dependent variables, coefficients, F-statistics, R-squared statistics, and alternative fixed-effect specifications. They do not report negative binomial estimates of incident counts. We describe their common structure below. The notation summarizes the model families represented in the tables; exact predictor timing, estimation commands, and inference procedures must be reconciled with the original code.

$$\text{CGP_it} = \alpha + \rho \text{CGP_i,t-1} + \beta \text{CSRP_it} + \gamma'X_it + \text{fixed effects} + \varepsilon_it. \quad (1)$$

$$\text{ESG_Con_it} = \alpha + \rho \text{ESG_Con_i,t-1} + \beta Z_it + \gamma'X_it + \text{fixed effects} + \varepsilon_it. \quad (2)$$

$$\text{ESG_Con_it} = \alpha + \rho \text{ ESG_Con_i,t-1} + \beta_1 \text{ CSRP_it} + \beta_2 \text{ M_it} + \beta_3 (\text{CSRP_it} \times \text{M_it}) + \gamma' \text{X_it} + \text{fixed effects} + \varepsilon_{\text{it}}. \quad (3)$$

In Equation (2), Z is governance or CSR performance. In Equation (3), M is governance, reporting, or assurance. X denotes the controls included in the relevant table column. Table 4 adds Asset, Lev, and Cash to the governance model. Tables 5 and 6 also include Accruals, MTB, and SaleGr in the controlled controversy models. The four specifications within each model family progress from a baseline model to controls, year and industry indicators, and year and firm indicators. These formulas include the lagged outcome shown in the tables and omit Age, which is absent from the reported results.

For continuous governance moderation, the marginal CSR association is $\beta_1 + \beta_3 \text{CGP}$. For reporting or assurance, it is β_1 for the zero-indicator group and $\beta_1 + \beta_3$ for the one-indicator group. The moderator's main coefficient evaluates the group difference when CSRP equals zero, outside the reported CSRP range. Accordingly, the main reporting or assurance coefficient is not an unconditional average effect. Marginal associations and uncertainty intervals at observed score values are needed for substantive interpretation.

Because scores are bounded and ESG_Con is concentrated at 100, linear models may not adequately represent the outcome distribution. Lagged dependent variables with firm effects also raise dynamic-panel concerns. Additional specifications, clustered inference, and sensitivity analysis are needed to evaluate these issues. These are proposed checks rather than additional results reported in this study.

4. Results

4.1. Sample composition and descriptive statistics

Table 1 indicates substantial variation in reporting and assurance across years and sectors. Observations increase from 151 in 2003 to 1,971 in 2017, before declining to 1,283 in 2019. The number of reports reaches 547 in 2018, and assured reports reach 172 that year. The earliest reporting and assurance proportions rely on small denominators and should be interpreted cautiously. Later changes in reporting incidence may reflect sample composition or coverage rather than a general retreat from sustainability disclosure.

Consumer Discretionary and Industrials account for 19.0% and 16.2% of observations, respectively. Reporting incidence is particularly high in Utilities, where 368 of 583 observations contain reports, and Consumer Staples, where 390 of 664 observations contain reports. In contrast, Financials has 265 reports among 1,398 observations. Consumer Staples also has 154 assured reports, or approximately 39.5% of its reports. These differences motivate accounting for industry composition rather than interpreting the pooled reporting indicator as a uniform practice.

Table 2 reports a mean ESG_Con of 87.5833, a median of 100, a minimum of 7.5096, and a standard deviation of 25.8879. The median at the upper bound indicates a substantial concentration of observations at 100. Under conventional coding, many observations have favorable recorded controversy performance, while a smaller group has substantially lower scores. This distribution is incompatible with describing ESG_Con as the number of incidents.

Governance and CSR performance vary widely. CGP has a mean of 50.0669 and standard deviation of 21.9089; CSRP has a mean of 36.3383 and standard deviation of 22.9065. Their ranges are 6.3532–88.4136 and 5.0350–82.7886, respectively. Accruals average 0.0043 and span both positive and negative values. The sign of working-capital accruals alone does not identify discretionary earnings management or fraudulent reporting.

The controls also exhibit substantial dispersion. Average leverage is 0.5677, with a maximum exceeding one. MTB includes negative values, while sales growth ranges from –0.4521 to 1.4997. These observations require checking variable definitions, denominators, and influential values. Interpretation also requires explicit documentation of winsorization or other outlier treatment.

4.2. Correlations

Table 3 reports Pearson correlations below the diagonal and Spearman correlations above it. Pearson correlations of ESG_Con with CGP, CSRP, CSR_Report, and CSR_Assurance are –0.144, –0.286, –0.160, and –0.157, respectively. The corresponding rank correlations are also negative. Under the conventional score direction, favorable governance and CSR measures are associated with less favorable controversy scores at the bivariate level. These correlations do not support a simple claim that sustainability practices prevent recorded controversies.

CGP and CSRP are positively correlated, with a Pearson coefficient of 0.453. Firm size is positively related to governance, CSR performance, reporting, and assurance. The Pearson correlation between CSRP and CSR_Report is 0.734, indicating substantial overlap between rated performance and disclosure participation. This association warrants model-specific collinearity diagnostics. A correlation threshold alone cannot establish that an interaction model is free of multicollinearity.

Most accrual correlations with governance and CSR measures are modestly negative. They are not evidence of fraud reduction because the reported accrual variable is not identified as a fraud indicator. Correlations describe unconditional relationships and cannot resolve common causes, coverage differences, or the endogenous choice to report and obtain assurance.

4.3. Governance performance

Table 4 reports positive CSRP coefficients in all four governance models: 0.0847, 0.0606, 0.0521, and 0.0987. Each is marked significant at the 1% level in the source

table. The association persists after adding controls, year and industry indicators, and year and firm indicators. A ten-point difference in CSRP corresponds to an estimated CGP difference of approximately 0.52–0.99 points across specifications, holding the included regressors constant. This pattern is consistent with H1.

Lagged CGP is positive throughout, with coefficients ranging from 0.3982 to 0.7680. Governance scores therefore show persistence in the reported models. The lower lag coefficient with firm indicators suggests that stable firm characteristics account for some pooled persistence. Firm size is positively associated with governance in the controlled models. Leverage is not statistically significant, and the cash coefficient changes from negative in pooled specifications to positive and marginally significant with firm indicators.

R-squared rises from 64.54% in the baseline model to 76.05% in the firm-indicator model. This increase should not be treated as causal validation, particularly because lagged outcomes and numerous indicators can explain substantial variation. The positive CSR association is compatible with complementarity, but reverse causality and shared measurement remain possible. The tables alone do not establish that environmental and social improvements cause later governance improvements; predictor timing and dynamic specification must first be verified.

4.4. Governance CSR performance and controversy scores

Table 5 presents three groups of four controversy-score models. In Columns 1–4, CGP coefficients are negative: -0.0789 , -0.0228 , -0.0263 , and -0.0059 . The second is significant at 10%, the third at 5%, and the first and fourth at 1%, as printed. In Columns 5–8, CSRP coefficients are also negative: -0.2125 , -0.0712 , -0.0754 , and -0.0791 . The first three are significant at 1% and the fourth at 5%. A ten-point CSRP difference corresponds to a -0.791 -point score difference in Column 8.

Under the conventional scoring direction, these negative associations imply less favorable recorded controversy performance at higher governance or CSR scores. They therefore contradict the simple prevention predictions in H2 and H3. If the research files establish that ESG_Con was transformed into an increasing exposure measure, the substantive interpretation would differ. The hypothesis assessment follows the documented score convention.

The interaction between CSRP and CGP is negative in Columns 9–12, with coefficients of -0.0016 , -0.0022 , -0.0021 , and -0.0010 . The first and fourth are marked significant at 1%, and the middle two at 5%. A negative interaction means that the CSR–score slope becomes more negative as governance increases. Under conventional score coding, it does not support the claim that governance strengthens controversy prevention in H4.

Interaction-model main coefficients must be interpreted conditionally. In Column 12, the CSRP coefficient is 0.0255 and the CGP coefficient is 0.0284. Using the rounded interaction coefficient and sample means, the CSR slope at mean CGP is

approximately $0.0255 - 0.0010 \times 50.0669 = -0.0246$. The governance slope at mean CSR is approximately $0.0284 - 0.0010 \times 36.3383 = -0.0079$. These are point-estimate illustrations derived from the printed values. Their significance cannot be determined without the coefficient covariance matrix.

The lagged outcome is generally positive and significant. Column 8 reports a notably small lag coefficient, 0.0219, compared with the 0.4680 coefficient in the firm-indicator interaction model in Column 12. This discrepancy should be checked against the original estimation output. R-squared ranges from 24.01% to 51.65%, but model fit does not resolve the direction or meaning of the controversy score. The reported Durbin–Watson statistics are not a substitute for inference appropriate to a dynamic panel.

4.5. Reporting and assurance

Table 6 first examines reporting in the full sample. The $\text{CSR} \times \text{CSR_Report}$ coefficients are -0.1289 , -0.2262 , -0.1988 , and -0.1290 , each marked significant at 1%. After controls and indicators are added, the standalone CSR coefficients are close to zero and not marked significant. The CSR slope therefore becomes more negative among reporters than among nonreporters. With the conventional score direction, this pattern does not support the prevention prediction in H5.

Column 4 illustrates the conditional interpretation. The CSR slope is 0.0027 for nonreporters and approximately -0.1263 for reporters, using the printed coefficients. The reporting-group contrast at mean CSR is approximately $5.0751 - 0.1290 \times 36.3383 = 0.3875$ score points. The positive main reporting coefficient is consequently not an unconditional reporting benefit; it is evaluated at CSR equal to zero. Neither the group contrast nor either combined slope has a recoverable standard error from the table alone.

Columns 5–8 examine assurance in the full sample. The assurance main coefficients are not marked significant. The interaction coefficients are negative, but statistical support varies: -0.0836 is unmarked, -0.0644 is significant at 10%, -0.0739 at 5%, and -0.2021 at 10%. The comparison group includes both nonreporters and reporters without assurance, limiting its interpretation as a comparison of assurance practices.

Columns 9–12 restrict the sample to 3,952 reporters. Their assurance interaction coefficients are -0.0909 , -0.0909 , -0.0813 , and -0.0629 . The first three are marked significant at 5% and the fourth at 1%. This pattern indicates a more negative CSR–score slope among assured reporters. Under conventional coding, it is not evidence that assurance reduces controversy exposure. In Column 12, the implied CSR slopes are 0.0373 for unassured reporters and approximately -0.0256 for assured reporters. Their individual significance cannot be inferred from the significance of the interaction alone.

The standalone assurance coefficients within reporters are positive, with weak significance in Columns 9 and 10 and no significance markers in Columns 11 and

12. These coefficients are conditional on CSRP equal to zero and therefore should not be described as average assurance effects. The full-sample mean of CSRP is not necessarily the reporter-only mean; interpreting reporter-group contrasts requires the subsample distribution and covariance estimates.

4.6. Overall interpretation

The strongest consistent finding is the positive relationship between environmental and social performance and governance scores. The controversy models show negative score associations and negative interactions, but their substantive meaning differs from the prevention hypotheses. A conventional provider-score interpretation suggests a disconnect between favorable rated practices and recorded controversy performance. Selection, scrutiny, disclosure visibility, and underlying risk could each contribute to this pattern.

The tables do not distinguish these mechanisms. They also do not establish that reports are misleading, that firms conceal financial reporting fraud, or that assurance increases or decreases fraud. Those conclusions would require direct measures of misrepresentation, intentional conduct, financial reporting violations, or independently verified incident outcomes. The evidence is best presented as a study of sustainability credibility and observed controversy scores, subject to the documented measurement and design limitations.

Conclusion

This study examines environmental and social performance, governance, reporting, assurance, and ESG controversy scores using the 12,295 firm-year observations reported for 2003–2019. Environmental and social performance is positively associated with governance scores across the reported models. Governance and CSR performance have negative associations with controversy scores, and governance, reporting, and assurance interactions further condition the CSR–score relationship. Under the conventional provider scoring direction, these negative coefficients do not indicate fewer controversies. H1 is consistent with the reported evidence; the prevention predictions in H2–H5 are not supported under that convention. Confirming the actual historical data and transformations is essential before finalizing these interpretations. The evidence establishes associations rather than causality, and it does not directly identify fraudulent conduct.

The broader contribution is to distinguish sustainability ratings and communication from independently observable corporate behavior. Credibility requires attention to measurement direction, underlying evidence, reporting and assurance scope, and the mechanisms through which adverse conduct becomes visible. Research and practice can advance by testing these relationships with clearer constructs, reproducible data procedures, and outcomes that separate misconduct, detection, disclosure, and remediation.

The findings and measurement issues motivate separating rated sustainability practices from observed conduct. Future studies should use explicit definitions of environmental performance, social performance, governance capability, reporting, assurance, and controversy exposure. Constructing CSR performance without governance reduces circularity in governance tests, but common-provider scoring and overlapping inputs still require examination. Provider disagreement and coverage differences should be treated as features of the measurement process rather than incidental noise. A useful extension would distinguish prevention from detection and public visibility. One outcome could capture independently verified adverse events, another their disclosure or media coverage, and a third their inclusion in a commercial database. Such a design would help determine whether high-scoring firms experience more underlying misconduct or whether their incidents are more likely to become observable. Severity, duration, recurrence, and remediation should be distinguished from the number of events or a composite score.

Boards and managers should assess sustainability commitments alongside operating controls, reliable performance information, stakeholder complaints, and corrective action. A high score or an extensive report should prompt questions about the underlying evidence and reporting boundaries. Governance structures need resources, competence, and authority to challenge management and address material weaknesses. The reported associations do not demonstrate the effectiveness of any particular control, but they underscore the limits of relying on visible structures alone.

Investors and analysts should distinguish a firm's performance ratings from its controversy history. They should verify whether a score increases with favorable performance or adverse exposure and whether the measure incorporates media attention, severity adjustments, or provider-specific classifications. Comparisons across firms and providers require attention to coverage, timing, and materiality. Reporting and assurance indicators are useful starting points, but their presence cannot substitute for assessment of content and scope.

Assurance practitioners should make engagement criteria, scope, limitations, and conclusions understandable to intended users. Evidence supporting favorable disclosures should be considered alongside contradictory information and the treatment of material incidents. Forensic accountants can contribute by reconstructing events, tracing source records, evaluating inconsistencies, and investigating alleged misrepresentation. Their role in an investigation is distinct from a general assertion that ESG-centric firms commit fraud.

Regulators and standard setters can support comparability through transparent definitions, reporting boundaries, and assurance disclosures. The study does not establish that a mandate would reduce misconduct or that a particular assurance regime is superior. Its practical implication is narrower: sustainability credibility depends on how claims are measured, supported, challenged, and connected to observable outcomes.

Several limitations restrict interpretation. First, the exact historical score definitions and any transformation must be verified against the extracted data. Public methodology establishes the conventional field direction, but it does not prove how the manuscript's variable was coded. Conclusions about H2–H5 depend on this verification. Geographic scope, number of unique firms, sample filters, matching procedures, and database extraction dates also remain insufficiently documented in the available research documentation. Second, the reported linear models use bounded outcomes and lagged dependent variables. The concentration of controversy scores at 100 may require specifications that distinguish the presence of controversies from score variation among affected firms. Firm effects and lagged outcomes raise dynamic-panel bias concerns. The standard-error estimator and any clustering are not identified, and the reported significance markers have been retained without re-estimation. The same printed observation count across pooled and lagged full-sample models also warrants reconciliation with missing-data and lag construction rules.

Third, observational associations do not establish causal direction. Governance may affect CSR, CSR may affect governance, and common organizational or market conditions may affect both. Reporting and assurance are endogenous choices. Year and firm indicators remove some variation but do not resolve time-varying omitted variables or selection. Media visibility and database coverage may influence recorded controversies independently of underlying conduct. Finally, this analysis has no direct fraud outcome. The findings cannot determine whether firms intentionally misstate sustainability information, use reporting to conceal financial fraud, or achieve lower-cost financing through deception. Future studies can address these questions using independently verified outcomes and a design that connects reporting claims to specific conduct. The present evidence supports a more limited discussion of sustainability practices, disclosure credibility, and observed controversies.

References

- Albuquerque, R., Koskinen, Y., Zhang, C (2019). Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science* 65(10), 4451–4469. <https://doi.org/10.1287/mnsc.2018.3043>
- Aouadi, A., Marsat, S (2018). Do ESG controversies matter for firm value? Evidence from international data. *Journal of Business Ethics* 151(4), 1027–1047. <https://doi.org/10.1007/s10551-016-3213-8>
- Barkemeyer, R., Revelli, C., Douaud, A (2023). Selection bias in ESG controversies as a risk for sustainable investors. *Journal of Cleaner Production* 405, 137035. <https://doi.org/10.1016/j.jclepro.2023.137035>

- Barnett, M.L (2007). Stakeholder influence capacity and the variability of financial returns to corporate social responsibility. *Academy of Management Review* 32(3), 794–816. <https://doi.org/10.5465/amr.2007.25275520>
- Berg, F., Kölbel, J.F., Rigobon, R (2022). Aggregate confusion: The divergence of ESG ratings. *Review of Finance* 26(6), 1315–1344. <https://doi.org/10.1093/rof/rfac033>
- Busch, T., Friede, G (2018). The robustness of the corporate social and financial performance relation: A second-order meta-analysis. *Corporate Social Responsibility and Environmental Management* 25(4), 583–608. <https://doi.org/10.1002/csr.1480>
- Christensen, D.M., Serafeim, G., Sikochi, A (2022). Why is corporate virtue in the eye of the beholder? The case of ESG ratings. *The Accounting Review* 97(1), 147–175. <https://doi.org/10.2308/TAR-2019-0506>
- Christensen, H. B., Hail, L., and Leuz, C (2021). Mandatory CSR and sustainability reporting: Economic analysis and literature review. *Review of Accounting Studies*, 26(3), 1176–1248. <https://doi.org/10.1007/s11142-021-09609-5>
- Cohen, J. R., Holder-Webb, L., and Zamora, V. L (2015). Nonfinancial information preferences of professional investors. *Behavioral Research in Accounting*, 27(2), 127–153. <https://doi.org/10.2308/bria-51185>
- Cuadrado-Ballesteros, B., Martínez-Ferrero, J., García-Sánchez, I.M (2017). Mitigating information asymmetry through sustainability assurance: The role of accountants and levels of assurance. *International Business Review* 26(6), 1141–1156. <https://doi.org/10.1016/j.ibusrev.2017.04.009>
- Delmas, M.A., Burbano, V.C (2011). The drivers of greenwashing. *California Management Review*, 54(1), 64–87. <https://doi.org/10.1525/cmr.2011.54.1.64>
- Donaldson, T., Preston, L.E (1995). The stakeholder theory of the corporation: Concepts, evidence, and implications. *Academy of Management Review*, 20(1), 65–91. <https://doi.org/10.5465/amr.1995.9503271992>
- Dowling, J., Pfeffer, J (1975). Organizational legitimacy: Social values and organizational behavior. *Pacific Sociological Review* 18(1), 122–136. <https://doi.org/10.2307/1388226>
- Eccles, R.G., Ioannou, I., Serafeim, G (2014). The impact of corporate sustainability on organizational processes and performance. *Management Science*, 60(11), 2835–2857. <https://doi.org/10.1287/mnsc.2014.1984>
- 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>

- Godfrey, P.C., Merrill, C.B., Hansen, J.M (2009). The relationship between corporate social responsibility and shareholder value: An empirical test of the risk-management hypothesis. *Strategic Management Journal*, 30(4), 425–445. <https://doi.org/10.1002/smj.750>
- Huang, X. B., and L. Watson (2015). Corporate social responsibility research in accounting. *Journal of Accounting Literature*, 34, 1–16. <https://doi.org/10.1016/j.acclit.2015.03.001>
- Khan, M., Serafeim, G. and Yoon, A (2016). Corporate sustainability: First evidence on materiality. *The Accounting Review*, 91(6), 1697–1724. <https://doi.org/10.2308/accr-51383>
- LSEG. (2022). Environmental, social and governance scores from LSEG. May. https://www.lseg.com/content/dam/ftse-russell/en_us/documents/other/lseg-esg-scores-methodology.pdf
- Lys, T., Naughton, J., Wang, C (2015). Signaling through corporate accountability reporting. *Journal of Accounting and Economics*, 60(1), 56-72. <https://doi.org/10.1016/j.jacceco.2015.03.001>
- Marquis, C., Toffel, M.W., Zhou, Y (2016). Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organization Science*, 27(2), 483–504. <https://doi.org/10.1287/orsc.2015.1039>
- O'Dwyer, B., Owen, D., and Unerman, J (2011). Seeking legitimacy for new assurance forms: The case of assurance on sustainability reporting. *Accounting, Organizations and Society*, 36(1), 31-52. <https://doi.org/10.1016/j.aos.2011.01.002>
- O'Dwyer, B., Owen, D.L (2005). Assurance statement practice in environmental, social and sustainability reporting: A critical evaluation. *The British Accounting Review*, 37(2), 205–229. <https://doi.org/10.1016/j.bar.2005.01.005>
- Rezaee, Z (2016). Business sustainability research: A theoretical and integrated perspective. *Journal of Accounting Literature*, 36, 48–64. <https://doi.org/10.1016/j.acclit.2016.05.003>
- Servaes, H., Tamayo, A (2013). The impact of corporate social responsibility on firm value: The role of customer awareness. *Management Science*, 59(5), 1045–1061. <https://doi.org/10.1287/mnsc.1120.1630>
- Simnett, R., Vanstraelen, A., Chua, W.F (2009). Assurance on sustainability reports: An international comparison. *The Accounting Review*, 84(3), 937–967. <https://doi.org/10.2308/accr.2009.84.3.937>
- Venter, E. R., and van Eck, L (2021). Research on extended external reporting assurance: Trends, themes, and opportunities. *Journal of International Financial Management and Accounting*, 32(1), 63-103. <https://doi.org/10.1111/jifm.12125>

Appendix

Table 1: Sample distribution

Panel A Distribution by year

Year	Obs.	Sample %	Reports	Reports %	Assured reports	Assured %
2003	151	1.2%	6	0.2%	4	0.4%
2004	154	1.3%	9	0.2%	4	0.4%
2005	255	2.1%	16	0.4%	5	0.5%
2006	283	2.3%	23	0.6%	9	0.9%
2007	283	2.3%	95	2.4%	18	1.7%
2008	310	2.5%	116	2.9%	22	2.1%
2009	388	3.2%	148	3.7%	25	2.4%
2010	524	4.3%	210	5.3%	41	3.9%
2011	567	4.6%	249	6.3%	54	5.1%
2012	588	4.8%	273	6.9%	75	7.1%
2013	670	5.4%	308	7.8%	91	8.6%
2014	778	6.3%	351	8.9%	98	9.3%
2015	811	6.6%	369	9.3%	109	10.3%
2016	1371	11.2%	446	11.3%	123	11.6%
2017	1971	16.0%	505	12.8%	145	13.7%
2018	1908	15.5%	547	13.8%	172	16.3%
2019	1283	10.4%	281	7.1%	61	5.8%
Total	12295	100.0%	3952	100.0%	1056	100.0%

Note: Percentages use the column total as the denominator. Reporting and assurance incidence discussed in the text instead use observations or reports as the denominator. Counts and percentages are retained from the supplied table.

Table 1: Sample distribution**Panel B.** Distribution by sector

Sector	Obs.	Sample %	Reports	Reports %	Assured reports	Assured %
Health Care	1403	11.4%	339	8.6%	93	8.8%
Financials	1398	11.4%	265	6.7%	71	6.7%
Energy	812	6.6%	292	7.4%	57	5.4%
Consumer Discretionary	2335	19.0%	644	16.3%	143	13.5%
Basic Materials	522	4.2%	263	6.7%	73	6.9%
Consumer Staples	664	5.4%	390	9.9%	154	14.6%
Technology	1206	9.8%	360	9.1%	102	9.7%
Telecommunications	414	3.4%	109	2.8%	39	3.7%
Industrials	1992	16.2%	658	16.6%	150	14.2%
Utilities	583	4.7%	368	9.3%	83	7.9%
Real Estate	966	7.9%	264	6.7%	91	8.6%
Total	12295	100.0%	3952	100.0%	1056	100.0%

Note: Percentages use the column total as the denominator. Reporting and assurance incidence discussed in the text instead use observations or reports as the denominator. Counts and percentages are retained from the supplied table.

Table 2: Descriptive statistics

Variable	Mean	Median	Maximum	Minimum	Std. Dev.
ESG_Con	87.5833	100.0000	100.0000	7.5096	25.8879
CGP	50.0669	50.6400	88.4136	6.3532	21.9089
CSRP	36.3383	30.5250	82.7886	5.0350	22.9065
Accruals	0.0043	0.0023	0.2178	-0.1944	0.0604
Asset	15.3061	15.2286	18.5803	10.6991	1.5375
Cash	0.1014	0.0715	0.6296	0.0007	0.1024
Lev	0.5677	0.5711	1.1775	0.0797	0.2275
MTB	3.6334	2.6700	18.9620	-4.5690	3.9329
SaleGr	0.0835	0.0549	1.4997	-0.4521	0.2271

Note: ESG_Con is the reported Asset4 controversy score (TRESGCS). CGP is governance performance (CGSCORE). CSRP averages environmental (ENSCORE) and social (SOSCORE) scores. Asset is log total assets; Cash is cash/assets; Lev is debt/assets; MTB is the market-to-book ratio of assets; SaleGr is sales growth. Accruals is the working-capital-change measure described in Section 4; scaling and adjustments require confirmation. The controversy-score interpretation follows conventional provider coding, subject to verification of the extracted variable.

Table 3: Correlation matrix

Variable	ESG_Con	CGP	CSRP	Report	Assurance	Accruals	Asset	Cash	Lev	MTB	SaleGr
ESG_Con	1	-0.145***	-0.338***	-0.174***	-0.179***	0.037*	-0.49***	-0.012	-0.048**	-0.057***	0.06***
CGP	-0.144***	1	0.415***	0.376***	0.281***	-0.049**	0.251***	-0.06***	0.068***	-0.064***	-0.198***
CSRP	-0.286***	0.453***	1	0.638***	0.447***	-0.102***	0.463***	0.086***	0.106***	0.084***	-0.257***
Report	-0.16***	0.4***	0.734***	1	0.423***	-0.108***	0.267***	0.033	0.179***	-0.047**	-0.303***
Assurance	-0.157***	0.281***	0.464***	0.423***	1	-0.099***	0.331***	0.027	0.117***	0.036*	-0.118***
ACCRUALS	0.021	-0.047**	-0.115***	-0.114***	-0.082***	1	-0.073***	-0.057***	-0.14***	0.003	0.169***
ASSET	-0.452***	0.255***	0.46***	0.259***	0.324***	-0.057***	1	-0.194***	0.232***	-0.075***	-0.116***
CASH	0.027	-0.051**	0.056**	0.014	-0.004	-0.147***	-0.211***	1	-0.168***	0.249***	0.034
LEV	-0.047**	0.075***	0.134***	0.186***	0.129***	-0.132***	0.221***	-0.126***	1	0.038*	-0.223***
MTB	-0.035	-0.021	0.059***	-0.006	0.084***	-0.029	-0.03	0.143***	0.111***	1	0.143***
SALEGR	0.059***	-0.214***	-0.253***	-0.27***	-0.097***	0.126***	-0.104***	0.031	-0.176***	0.054**	1

Note: Pearson correlations appear below the diagonal and Spearman correlations above it. Report denotes CSR_Report and Assurance denotes CSR_Assurance. Significance markers are retained as printed; the source correlation table does not identify its inference procedure.

Table 4: Governance performance and CSR performance

Variable	Model 1	Model 2	Model 3	Model 4
Lagged CGP	0.7599*** (143.1304)	0.768*** (131.7214)	0.7646*** (129.7853)	0.3982*** (40.1464)
CSRP	0.0847*** (14.8095)	0.0606*** (8.9122)	0.0521*** (7.1733)	0.0987*** (6.714)
Asset		0.3494*** (3.5424)	0.5199*** (4.6035)	1.2458*** (3.4038)
Lev		0.0164 (0.0307)	-0.1512 (-0.2754)	-0.4713 (-0.4083)
Cash		-3.853*** (-3.5948)	-2.2959** (-2.0146)	3.2945* (1.8381)
Constant	10.1379*** (40.6626)	5.6419*** (3.9958)	1.7244 (1.0054)	9.4389 (1.6197)
Year indicators	No	No	Yes	Yes
Industry indicators	No	No	Yes	No
Firm indicators	No	No	No	Yes
R ²	64.54%	66.36%	66.81%	76.05%
Adjusted R ²	64.54%	66.35%	66.73%	71.22%
F	13853.9000	4849.4090	796.3788	15.7304
F-test p-value	0.0001	0.0001	0.0001	0.0001
DW	2.3152	2.2652	2.2721	2.1647
Obs.	12295	12295	12295	12295

Note: Dependent variable: CGP. Parenthetical statistics and significance markers are retained from the supplied table. *** $p < .01$; ** $p < .05$; * $p < .10$. DW denotes the reported Durbin–Watson statistic. Predictor timing and the standard-error estimator require verification.

Table 5: Governance CSR performance and controversy scores**Panel A.** Governance performance

Variable	Model 1	Model 2	Model 3	Model 4
Lagged ESG_Con	0.495*** (44.5051)	0.3848*** (30.725)	0.3672*** (29.9968)	0.4813*** (42.1717)
CGP	-0.0789*** (-9.2438)	-0.0228* (-1.6639)	-0.0263** (-2.0085)	-0.0059*** (-2.8104)
CSRP				
CSRP × CGP				
Accruals		2.2746 (0.4956)	6.0366 (1.3184)	0.9282* (1.7072)
Asset		-4.79*** (-20.7146)	-5.0261*** (-21.9427)	-1.1954*** (-17.0718)
Cash		-12.0191*** (-4.1392)	-8.8459*** (-3.076)	-3.2879*** (-10.1783)
Lev		2.5664* (1.8952)	1.7438 (1.3644)	0.5681*** (2.6577)
MTB		-0.1693** (-2.4179)	-0.1389** (-2.0195)	-0.0612*** (-4.3531)
SaleGr		-2.1865* (-1.9026)	-1.0317 (-0.8817)	-0.3325* (-1.9333)
Constant	47.4702*** (40.0895)	128.134*** (32.3557)	134.8881*** (27.2553)	68.4776*** (41.2054)
Year indicators	No	No	Yes	Yes
Industry indicators	No	No	Yes	No
Firm indicators	No	No	No	Yes
R ²	24.01%	29.16%	31.50%	34.08%
Adjusted R ²	24.00%	29.08%	31.20%	33.88%
F	2404.9140	366.5039	105.3017	175.0098
F-test p-value	0.0001	0.0001	0.0001	0.0001
DW	2.3590	2.1874	2.1711	2.2564
Obs.	12295	12295	12295	12295

Note: Dependent variable: ESG_Con. Numerical results are retained without re-estimation. Parenthetical statistics and significance markers are reproduced as supplied. *** $p < .01$; ** $p < .05$; * $p < .10$. The field convention assigns higher scores to more favorable controversy performance; any transformation must be confirmed. Blank cells indicate regressors omitted in that model.

Table 5: Governance CSR performance and controversy scores**Panel B.** CSR performance

Variable	Model 5	Model 6	Model 7	Model 8
Lagged ESG_Con	0.4458*** (54.6803)	0.378*** (31.0552)	0.3608*** (29.4349)	0.0219* (1.6557)
CGP				
CSRP	-0.2125*** (-23.6261)	-0.0712*** (-4.6571)	-0.0754*** (-4.8022)	-0.0791** (-2.4596)
CSRP × CGP				
Accruals		0.1861 (0.0404)	3.9525 (0.8637)	6.4023 (1.4388)
Asset		-4.2182*** (-16.9806)	-4.414*** (-16.5504)	-6.7459*** (-7.2538)
Cash		-10.5545*** (-3.7696)	-7.2791** (-2.5354)	6.0755 (1.2741)
Lev		2.7786** (2.2182)	1.6808 (1.3229)	-4.1953 (-1.4695)
MTB		-0.1604** (-2.3357)	-0.1245* (-1.8054)	0.1005 (1.1079)
SaleGr		1.2093 (0.9982)	2.0842* (1.7004)	1.7453 (1.348)
Constant	54.7322*** (61.8715)	120.8236*** (29.8059)	125.3852*** (23.7283)	193.3378*** (13.3228)
Year indicators	No	No	Yes	Yes
Industry indicators	No	No	Yes	No
Firm indicators	No	No	No	Yes
R ²	26.32%	29.36%	31.76%	51.65%
Adjusted R ²	26.31%	29.28%	31.45%	40.64%
F	2718.4880	370.9779	103.4970	4.6897
F-test p-value	0.0001	0.0001	0.0001	0.0001
DW	2.3053	2.1806	2.1652	2.1503
Obs.	12295	12295	12295	12295

Note: Dependent variable: ESG_Con. Numerical results are retained without re-estimation. Parenthetical statistics and significance markers are reproduced as supplied. *** $p < .01$; ** $p < .05$; * $p < .10$. The field convention assigns higher scores to more favorable controversy performance; any transformation must be confirmed. Blank cells indicate regressors omitted in that model.

Table 5: Governance CSR performance and controversy scores**Panel C.** Governance interaction

Variable	Model 9	Model 10	Model 11	Model 12
Lagged ESG_Con	0.4448*** (54.5208)	0.3763*** (17.1929)	0.3588*** (16.3969)	0.468*** (13.9332)
CGP	0.047*** (3.1485)	0.0799*** (2.7272)	0.0708** (2.5592)	0.0284*** (4.3284)
CSRP	-0.1229*** (-4.9586)	0.0524 (0.9854)	0.0444 (0.8586)	0.0255** (2.3382)
CSRP × CGP	-0.0016*** (-3.9308)	-0.0022** (-2.4196)	-0.0021** (-2.4039)	-0.001*** (-3.8626)
Accruals		0.5459 (0.2148)	4.3144 (1.3078)	0.0331 (0.1687)
Asset		-4.1824*** (-8.0072)	-4.3758*** (-8.8386)	-0.8708*** (-4.9446)
Cash		-10.4033*** (-4.8488)	-7.1828*** (-2.7575)	-2.4415*** (-4.8548)
Lev		2.6642* (1.954)	1.5652 (1.4576)	0.3008*** (3.4634)
MTB		-0.1634** (-2.4603)	-0.1276** (-2.2132)	-0.0391*** (-3.3445)
SaleGr		1.391 (0.893)	2.2226* (1.9225)	0.0848* (1.7329)
Constant	52.4884*** (47.8335)	116.5901*** (14.2024)	121.9401*** (16.8834)	64.1139*** (13.6603)
Year indicators	No	No	Yes	Yes
Industry indicators	No	No	Yes	No
Firm indicators	No	No	No	Yes
R ²	26.39%	29.52%	31.90%	33.38%
Adjusted R ²	26.37%	29.42%	31.58%	33.15%
F	1364.4740	299.0112	98.0332	148.7232
F-test p-value	0.0001	0.0001	0.0001	0.0001
DW	2.3043	2.1780	2.1618	2.2470
Obs.	12295	12295	12295	12295

Note: Dependent variable: ESG_Con. Numerical results are retained without re-estimation. Parenthetical statistics and significance markers are reproduced as supplied. *** $p < .01$; ** $p < .05$; * $p < .10$. The field convention assigns higher scores to more favorable controversy performance; any transformation must be confirmed. Blank cells indicate regressors omitted in that model.

Table 6: Reporting assurance and controversy scores**Panel A.** Reporting in the full sample

Variable	Model 1	Model 2	Model 3	Model 4
Lagged ESG_Con	0.4436*** (21.5)	0.3694*** (16.8873)	0.3544*** (21.2304)	0.4611*** (34.1148)
CSRP	-0.1644*** (-5.9875)	-0.002 (-0.0664)	-0.0104 (-0.4202)	0.0027 (1.3602)
CSR_Report	5.6666*** (5.1977)	11.8518*** (11.4829)	9.856*** (5.5167)	5.0751*** (13.7019)
CSR_Assurance				
CSRP × CSR_Report	-0.1289*** (-5.7442)	-0.2262*** (-8.9786)	-0.1988*** (-5.3238)	-0.129*** (-12.6015)
CSRP × CSR_Assurance				
Accruals		0.3349 (0.1316)	5.7844 (1.2863)	1.1056*** (4.8445)
Asset		-4.2907*** (-8.1584)	-4.4269*** (-15.4592)	-0.8451*** (-18.1519)
Cash		-10.0402*** (-4.5785)	-6.9318** (-2.5632)	-1.9465*** (-13.8413)
Lev		2.3181* (1.6878)	1.1851 (0.9298)	0.3236*** (3.3796)
MTB		-0.1485** (-2.2465)	-0.0995 (-1.3177)	-0.0367*** (-5.3027)
SaleGr		1.4352 (0.9261)	-1.2327 (-1.0957)	-0.3403*** (-5.7975)
Constant	53.9*** (23.4576)	120.8762*** (14.6135)	125.5953*** (23.7751)	65.1111*** (43.0981)
Year indicators	No	No	Yes	Yes
Industry indicators	No	No	Yes	No
Firm indicators	No	No	No	Yes
R ²	26.44%	29.83%	32.06%	34.18%
Adjusted R ²	26.42%	29.73%	31.74%	33.96%
F	1367.4330	303.4465	98.5171	153.8083
F-test p-value	0.0001	0.0001	0.0001	0.0001
DW	2.3032	2.1738	2.1555	2.2462
Obs.	12295	12295	12295	12295

Note: Dependent variable: ESG_Con. Numerical results are retained without re-estimation. Parenthetical statistics and significance markers are reproduced as supplied. *** $p < .01$; ** $p < .05$; * $p < .10$. The field convention assigns higher scores to more favorable controversy performance; any transformation must be confirmed. Blank cells indicate regressors omitted in that model.

Table 6: Reporting assurance and controversy scores**Panel B.** Assurance in the full sample

Variable	Model 5	Model 6	Model 7	Model 8
Lagged ESG_Con	0.4886*** (29.5331)	0.4064*** (15.1503)	0.3795*** (12.4249)	0.0748*** (2.5831)
CSRP	-0.1788*** (-9.3668)	-0.0282 (-0.78)	-0.0031 (-0.075)	-0.0522 (-0.7579)
CSR_Report				
CSR_Assurance	4.9914 (1.1563)	4.0658 (1.5688)	3.1459 (1.4277)	10.7504 (1.3031)
CSRP × CSR_Report				
CSRP × CSR_Assurance	-0.0836 (-1.5808)	-0.0644* (-1.6832)	-0.0739** (-2.2754)	-0.2021* (-1.7127)
Accruals		-4.1762 (-0.3491)	1.3592 (0.1107)	-5.0477 (-0.4774)
Asset		-8.1333*** (-11.6709)	-8.7557*** (-10.2365)	-10.9093*** (-5.1524)
Cash		-9.5277 (-1.1206)	0.0219 (0.0019)	1.0528 (0.0787)
Lev		9.216** (2.5208)	6.2824 (1.4728)	-1.7508 (-0.2351)
MTB		-0.1549 (-0.825)	-0.0741 (-0.4095)	-0.057 (-0.2377)
SaleGr		-7.0505* (-1.832)	-6.1706 (-1.4858)	0.1319 (0.034)
Constant	47.829*** (26.6158)	176.1015*** (14.4311)	222.0324*** (12.9958)	262.5323*** (7.3662)
Year indicators	No	No	Yes	Yes
Industry indicators	No	No	Yes	No
Firm indicators	No	No	No	Yes
R ²	30.29%	36.89%	41.08%	56.60%
Adjusted R ²	30.23%	36.49%	39.81%	45.31%
F	479.9473	91.4242	32.5521	5.0151
F-test p-value	0.0001	0.0001	0.0001	0.0001
DW	2.2558	2.1831	2.1402	2.2324
Obs.	12295	12295	12295	12295

Note: Dependent variable: ESG_Con. Numerical results are retained without re-estimation. Parenthetical statistics and significance markers are reproduced as supplied. *** $p < .01$; ** $p < .05$; * $p < .10$. The field convention assigns higher scores to more favorable controversy performance; any transformation must be confirmed. Blank cells indicate regressors omitted in that model.

Table 6: Reporting assurance and controversy scores**Panel C.** Assurance among reporters

Variable	Model 9	Model 10	Model 11	Model 12
Lagged ESG_Con	0.569*** (26.9194)	0.4172*** (14.2994)	0.3824*** (12.656)	0.4826*** (27.7403)
CSRP	-0.0672* (-1.8576)	0.0576 (1.2023)	0.0222 (0.4417)	0.0373 (1.5409)
CSR_Report				
CSR_Assurance	4.482* (1.7991)	5.4546* (1.9284)	3.5018 (1.2429)	0.9279 (0.6474)
CSRP × CSR_Report				
CSRP × CSR_Assurance	-0.0909** (-2.4732)	-0.0909** (-2.1777)	-0.0813** (-2.051)	-0.0629*** (-2.6912)
Accruals		-21.909 (-1.2877)	-20.0134 (-1.2182)	-26.4464*** (-2.9069)
Asset		-9.5725*** (-11.8203)	-10.2584*** (-11.8744)	-7.8533*** (-16.6173)
Cash		-23.8459** (-2.3802)	-16.2909 (-1.4888)	-22.6496*** (-4.9892)
Lev		4.7152 (1.0066)	1.6479 (0.3133)	3.2414 (1.387)
MTB		-0.103 (-0.4495)	-0.0065 (-0.0277)	-0.071 (-0.5336)
SaleGr		-2.3001 (-0.4403)	-0.8609 (-0.1491)	-5.094* (-1.9557)
Constant	34.8835*** (12.1181)	198.507*** (13.5676)	207.8348*** (13.7752)	170.2203*** (21.568)
Year indicators	No	No	Yes	Yes
Industry indicators	No	No	Yes	No
Firm indicators	No	No	No	Yes
R ²	33.40%	39.73%	44.54%	73.30%
Adjusted R ²	33.26%	39.22%	43.05%	72.77%
F	233.8538	77.3252	29.8489	138.4818
F-test p-value	0.0001	0.0001	0.0001	0.0001
DW	2.2453	2.1927	2.1478	2.0313
Obs.	3952	3952	3952	3952

Note: Dependent variable: ESG_Con. Numerical results are retained without re-estimation. Parenthetical statistics and significance markers are reproduced as supplied. *** p < .01; ** p < .05; * p < .10. The field convention assigns higher scores to more favorable controversy performance; any transformation must be confirmed. Blank cells indicate regressors omitted in that model.

Dejan Malinić
University of Belgrade, Faculty of
Economics and Business
ORCID: 0000-0001-5525-1278

Nataša Džudović
LeadingMile d.o.o.

CONCEPTUAL FORENSIC FRAMEWORK FOR IDENTIFYING TRANSFER PRICING MANIPULATION RISK

UDK 336.2:657.6

DOI <https://doi.org/10.54318/czf.2026.dm.003>

Abstract

While transfer pricing as pricing of internal transfers of goods, services and capital within a group is a legitimate business practice, there have been many evidence of their misuse. Transfer pricing manipulation occurs when companies deliberately set these prices at non-market level in order to shift profits, usually to reduce the overall tax burden. This praxis not only erodes tax revenues but also distorts investment activities and can place domestic companies at a competitive disadvantage against multinationals with advanced tax planning strategies. In this paper we develop a conceptual forensic framework to identify the risk of transfer pricing manipulation and explore financial reporting capabilities and limitations to identify circumstances that enhance this risk. We group the identified risks into four dimensions of risk intensity: (1) risk of exposure to transactions with related parties (2) commercial transactions risk (3) financial transactions risk and (4) tax related risk. By employing a set of indicators we measure each dimension risk signals intensity and compute convergence risk that indicates signal activation. Based on obtained result we provide a risk matrix that assists with interpreting the results and identifying the risk of transfer pricing manipulation.

Key words: transfer pricing, manipulation, forensic framework, risk intensity, risk convergence, risk matrix

JEL Classification: M41; G30; H25; L25

Introduction

Rapid advancement of digitalisation and globalisation has profoundly reshaped societies, economies and international business operations worldwide. The growing number of multinational companies (MNC) have brought many challenges for the taxing authorities especially in cases when profits are being shifted to low or no-tax countries where little economic activities are being performed. The praxis of base erosion and profit shifting (BEPS) has had a global impact, resulting in estimated annual loss of 100-240 billion USD, which corresponds to 4-10% of global corporate income tax revenue.¹ Considering that a substantial proportion of global trade is performed by MNCs, transfer pricing, as pricing of transactions between related entities within group, has become one of the most significant mechanisms for profit allocations, opening opportunities for creative tax planning and transfer pricing manipulation.

The consequences of transfer pricing manipulation extend beyond the loss of tax revenue. Countries may experience reduced capacity to fund public services and infrastructure, while domestic companies operating solely within national borders may face competitive disadvantages compared to multinationals capable of engaging in sophisticated tax planning. Furthermore, profit shifting can undermine public confidence in the fairness and effectiveness of tax systems, particularly in developing economies where tax revenues present a critical source of government funding.

Given the growing significance of transfer pricing in international taxation, both tax policymakers and researchers have intensified efforts to understand the mechanisms through which transfer pricing manipulation occurs and develop measures to prevent abusive practices. International initiatives led by OECD have been focused on increasing tax transparency and aligning the taxation of profits with the location of economic activities and value creation. However, such efforts are yet to be fully realised. Transfer pricing manipulation remains a persistent challenge due to many difficulties associated with valuing unique assets and services, accessing international capital markets and corporate capital structures, as well as monitoring complex cross-border transactions. The challenges are more profound in developing economies.

In this paper we develop a conceptual forensic framework to identify the risk of transfer pricing manipulation. We examine financial reporting capabilities in assisting with identification of individual risks and assessment of overall exposure to transfer pricing manipulation risk. The paper is organised as follows. After introduction, we provide an overview of empirical evidence of transfer pricing manipulations, that is followed by a theoretical review of transfer pricing management challenges. In the next section, we provide and discuss a conceptual forensic framework at the core of which lies a risk matrix that can assist forensic

¹ <https://www.oecd.org/en/topics/base-erosion-and-profit-shifting-beps.html>

accountants in assessing and nuancing risk exposure to aggregated risk from transfer pricing manipulation.

1. The Evidence of Transfer Mispricing

As MNCs continue to expand across jurisdictions, taxation issues become increasingly complex for both tax authorities and companies. Their taxation can no longer be effectively addressed through national tax rules alone but instead requires coordination within international tax framework. The generally adopted OECD transfer pricing guidelines treat companies within the group as separate entities, an approach that taxes individual group members on the arm's length principle under which the intra-group transactions should be treated as if they occurred on the market and all effects of special conditions based on their internal relationships should be eliminated (OECD, 2022). This principle has been chosen to serve a dual objective. On one hand, it should secure the appropriate tax base in each jurisdiction. On the other hand, it should prevent double taxation that could distort the trade, undermine investments and create a conflict between tax administrations. If prices at which an entity transfers physical goods, intangible property or provides services or capital to associated entities do not reflect the market conditions, the tax liabilities of associated entities and tax revenues of host jurisdictions become disproportioned.

Manipulation of transfer pricing or transfer mispricing occurs when companies intentionally misprice their internal transactions to move profits away from high-tax into low-tax jurisdictions. By over-invoicing subsidiaries in low-tax countries and under-invoicing in high-tax countries, MNCs can artificially reallocate profits out of the countries with real economic activity in order to avoid taxation under higher rates. It has been estimated that MNCs shifted over 850 billion USD in profits in one year globally, primarily to countries with effective tax rates below 10% (Garcia-Bernardo & Janský, 2024). There are many empirical evidence identifying impacts of aggressive transfer pricing strategies.

On a sample of German MNCs, it was shown that 82% of multinationals subject to OECD country by country reporting requirement founded subsidiaries in tax havens that are notably more profitable than those in non-havens. Results suggested that approximately 40% of the profits reported in tax havens can be contributed to profit shifting activities. The associated annual tax base loss for Germany amounted to 5.4 billion EUR. Including estimated amounts for MNCs not covered by the OECD reporting requirement, the loss in tax base amounted to 19.1 billion EUR per year, corresponding to 4.3% of the total profits reported by German MNCs, i.e. 5.7 billion EUR in yearly tax revenue loss (Fuest, Hugger, & Neumeier, 2022).

Another study examining a decade of operations by Slovak companies with direct ownership to entities located in tax havens found that entities with such ownership reported weaker performance than entities without subsidiaries in these jurisdictions. Parent companies resident in countries with preferential tax regimes had a negative impact on financial performance of their subsidiaries, which exhibited lower

profitability as measured by return on assets, lower asset turnover, less favourable current liquidity ratios and a higher risk of bankruptcy. The study concluded that companies with ownership links to tax havens actively engage in profit shifting activities (Elexa, Ištók, & Šlampačková, 2022).

Next, subsidiaries of foreign MNCs were found to be more profitable than domestic companies in low-tax countries, while being less profitable than domestic companies in high-tax countries. Building on these profitability differentials, it was estimated that nearly 40% of MNCs' profits is being shifted to tax havens worldwide. Assuming all countries adopted the same effective corporate tax rate, while global profits and investment levels remained unchanged, it was estimated that reported profits would increase by approximately 15% in high-tax European Union countries and by 10% in the United States, while declining by around 60% in current tax haven jurisdictions (Tørsløv, Wier, & Zucman, 2023). Even current losses were affected by profit shifting incentives. Findings suggest that MNCs altered the distribution of reported profits to take advantage of losses and achieve smaller overall tax burden (De Simone, Klassen, & Seidman, 2017).

The result is further supported by empirical investigation of French MNCs that found deviations between intra-group transfer prices and prices determined under the arm's length principle to be particularly pronounced when affiliated companies were residents of tax havens (Davies, Martin, Parenti, & Toubal, 2018). Following research conducted in the same country found reducing profits shifting activities in response to more rigorous tax reporting requirements. However, this reduction in profit shifting was accompanied by a decline in investment activity in low-tax jurisdictions, whereas no comparable effect was found for entities operating in high-tax countries (Laudage Teles, et al., 2026).

Furthermore, evidence from Danish MNCs indicates that subsidiaries operating in low-tax jurisdictions reduce their exports by approximately 5.7% to 9.1% on average. This corresponds to an unreported yearly export value of approximately 141 million USD and a loss of tax revenue equivalent to 3.24% of the total reported profits of Danish MNCs (Cristea & Nguyen, 2016). Additionally, research showed that Nordic MNCs respond to varying corporate tax rates by shifting profits to countries with lower tax rates. The results demonstrated that a 10pp increase in the tax rate in the country where the profit was generated leads to a 7% to 13% reduction in reported pre-tax profit (Viertola, 2023). Next, research on export transactions conducted by MNCs in the United Kingdom revealed evidence of tax-motivated transfer pricing manipulation aimed at shifting profits to low-tax jurisdictions. The study found that transfer mispricing increases significantly under a worldwide taxation system compared to a territorial taxation approach. The authors also found that these effects are more pronounced when there are significant research and development activities and when transactions are concentrated in countries that are not classified as tax havens but instead impose low to medium corporate income tax rates (Liu, Schmidt-Eisenlohr, & Guo, 2020).

Illicit financial flows were also frequently examined, especially focusing on capital allocation and investment activity. Analysing a sample covering 27 countries over 12 year period, research showed a negative impact of transfer pricing regulations on MNCs investment levels. Compared to investments in similar fully domestic owned companies, investment in foreign affiliates was on average about 11% lower after introducing tax rules (De Mooij & Liu, 2020). Investment levels were also affected by introducing restrictions on foreign affiliates interest deductibility, commonly referred to as thin capitalisation rules. It was shown that the presence of typical thin capitalisation rules approximately doubles the sensitivity of foreign direct investment to corporate tax rates. In addition, it was identified that the introduction or tightening of such rules had significant negative effects on investment levels in high-tax countries. However, this analysis didn't find a significant relationship between transfer pricing regulations and the level of foreign direct investment (Buettner, Overesch, & Wamser, 2018). Many studies presented evidence of transfer pricing impacting capital structures of MNCs (Kalra & Afzal, 2023). A ten-year sample of German MNCs provided evidence of international internal financing being used when affiliated entities were exposed to lower tax rates. The size of such internal financing increased with rising difference between the host country tax rate and the lowest tax rate faced by the entities in the group. The study found that a 10pp increase in the host country tax rate increased the debt-to-equity ratio by about 2.1 pp (Buettner & Wamser, 2013). Next, based on European companies financial statements, research found that an increase in marginal effective tax rate by 10pp results in an increase in financial leverage by 2.59pp (Huizinga, Laeven, & Nicodème, 2008). Although offering a comprehensive overview, the results of previous research do not provide unambiguous conclusions about whether and to what extent internal debt is used to reallocate profits or whether these effects reflect the conventional understanding of financial leverage. Some found the effects of transfer pricing on financial leverage to be relatively small and not as exploited as a mechanism for profit shifting due to numerous regulatory restrictions on this type of financing (Fuest, Hebous, & Riedel, 2011).

While empirical evidence from developed economies demonstrates significant tax base erosion resulting from transfer pricing manipulation, developing countries experience magnified vulnerabilities and revenue losses (Moyo & Liu, 2022). A comprehensive study based on a sample of 210,000 corporations across 142 countries examined whether tax avoidance by MNCs is more prevalent in less developed economies. The findings indicated reported profits to be less responsive to profit shifting incentives with stronger economic and institutional development. This may help explain why many developing countries adopt relatively low corporate income tax rates despite their pressing needs for additional revenue and significant constraints they face in obtaining alternative sources of public revenue (Johannesen, Tørsløv, & Wier, 2020). Using panel data from 173 countries over a 33-year period, a similar study reached complementary conclusions, finding that international tax competition, measured by differences in corporate tax rates, contributes to tax base spill overs that are two to three times larger in developing

countries than in developed economies. Consequently, public revenue losses arising from economic activities linked to jurisdictions with preferential tax regimes are estimated to reduce government revenues by between one and two percentage points of GDP (Crivelli, De Mooij, & Keen, 2015). Despite the efforts of the legislator to enforce the arm's length principle, transfer pricing enforcement in developing countries has been severely constrained by legislative ambiguity, a lack of transparency and local comparability data, inadequate administrative capacity and flawed double taxation agreements (Eukeria & Mpofu, 2024).

In addition to above, one of big consequences of transfer pricing manipulation was identified in the decline in business dynamism, i.e. in decline in the number of entities that enter, expand, contract and exit the market. By altering market competition, capital allocation and compliance burdens across local and international environment, MNCs create tax advantages over local competitors (OECD, 2026). Though the magnitude of these effect remains unclear, it is evident that corporate income tax rate, tax compliance costs and tax incentives play a significant role in entities' ability to come to the market, especially for small companies.

Despite empirical research finding many evidence of transfer mispricing, some studies indicate that transfer prices are often used to meet tax legislation requirements rather than to exploit tax rules (Klassen, Lisowsky, & Mescall, 2017). Research showed profit shifting activities significantly reducing when tax legislation tightens requirements for transfer pricing documentation. Based on European high-tax countries data, stricter transfer pricing regulations were shown to increase the effective tax expenses of MNCs and reduce the volume of their internal transactions. The sales and profits of affected companies' domestic competitors rose significantly, while profit margins remained largely unchanged (Gauß, Kortenhaus, Riedel, & Simmler, 2024). Research on United States corporations showed that developing transfer pricing agreements with tax authorities reduced both the incentives and the capacity of MNCs to shift profits from high-tax to low-tax countries (Alghamdi, Eulaiwi, Hussein, Duong, & Taylor, 2024). Some studies even challenged the common view that transfer mispricing is more serious in developing countries than in advanced economies, and found that the OECD recommended reform hasn't had a lasting impact in detaining such practices (Wier, 2020). Contrary to this view, recent research showed that introducing a global minimum tax rate has already had an impact on relatively low-taxed MNCs that experienced an increase in their effective tax rates of 1.7pp, resulting in an increase of 79bn-109bn EUR in tax revenue globally in the first year of its implementation, which is equivalent to 2.4-3.4% of global corporate income tax revenue (Contreras, Hugger, & Zawisza, 2026).

It is clear that empirical research on transfer pricing presents a duality. While some studies emphasise its role in tax avoidance, others highlight compliance and their relationship with managerial objectives. In cases when mispricing in transaction between affiliated entities occurs, the consequences are shown through fiscal losses, distorted competitive markets, foreign direct investment and capital allocation that misrepresent actual financial performance and regional economic activity. This

underscores the necessity of examining the mechanisms of transfer pricing and developing a structured framework to evaluate its impacts on taxation, corporate behaviour and broader economic outcomes.

2. Theoretical Framework

Transfer pricing can carry significant economic, organisational and regulatory risk. The risks can be associated with large number of competing transfer pricing objectives, related to tax minimisation, performance management, resource allocation, but also regulatory compliance and stakeholder legitimacy. In an effort to explain why transfer pricing is inherently risky and challenging to manage, researchers have provided several theoretical perspectives.

The relationship between transfer pricing, profitability, tax aggressiveness and compliance can be explained by agency theory, suggesting a permanent conflict of interest between principals and agents. From transfer pricing perspective, principals can be shareholders but also central management of complex organisations, whereas the agents are managers on the lower end of the organisational structure (Jensen & Meckling, 1976). As shareholders and central management often lack relevant information, transfer pricing decisions may be delegated to organisational units, particularly in dynamic environments where processes cannot be fully formalised or standardised (Arnold, Elsinger, & Rankin, 2021). In such circumstances, the role of central management (the principal) is primarily to plan and allocate resources, monitor divisional performance and select and develop lower level managers (agents), who are responsible for achieving both divisional and organisational objectives. The theory suggests that the degree of delegation of transfer pricing decisions is determined by the level of information asymmetry and goal incongruence. Delegating decision rights to the agent with superior information enables the company to fully utilise agent's specific knowledge, but is more beneficial when the interests are closely aligned. A mismatch between organisational structure and the level of autonomy granted to organisational units can lead to negative outcomes, including lower satisfaction with internal transactions, lower perceived fairness and transfer pricing effectiveness among divisional managers.

Unlike agency theory that focuses on conflicts of interest, transfer pricing can be also explained through organisational control theory that focuses on the difficulty of pursuing multiple objectives simultaneously. The core principle states that there is no single transfer pricing that can simultaneously optimise tax compliance, managerial motivation, performance measurement, divisional autonomy and resource allocation (Eccles, 1985). Therefore, transfer pricing creates multiple risks, such as goal incongruence where a transfer price beneficial for central management may be harmful for divisional management, distortion of divisional profits and performance evaluations, internal conflicts between organisational units, disputes with tax authorities and more. The theory is considered to be one of the strongest

frameworks as transfer pricing has first originated as a management control and accounting issue, long before becoming a tax compliance problem.

On the other hand, institutional theory provides a very different explanation of transfer pricing associated risks and argues that organisations adopt transfer pricing practices due to the pressure to conform to societal, regulatory and professional expectations in order to maintain shareholder legitimacy and avoid being perceived as avoidant tax payers. Institutional pressures were identified in coercive pressures that derive from political influence and a legitimacy problem, in mimetic pressures resulting from companies imitating other companies when faced with uncertainty and in normative pressures deriving from professional standards and expectations (DiMaggio & Powell, 1983). The theory explains that transfer pricing decisions are not based solely on efficiency and profit maximisation objectives, but rather respond to regulatory requirements and professional norms, especially related to OECD standards, public scrutiny and investor pressures. Failure to meet these expectations may result in regulatory sanctions, reputational damage and increased compliance costs.

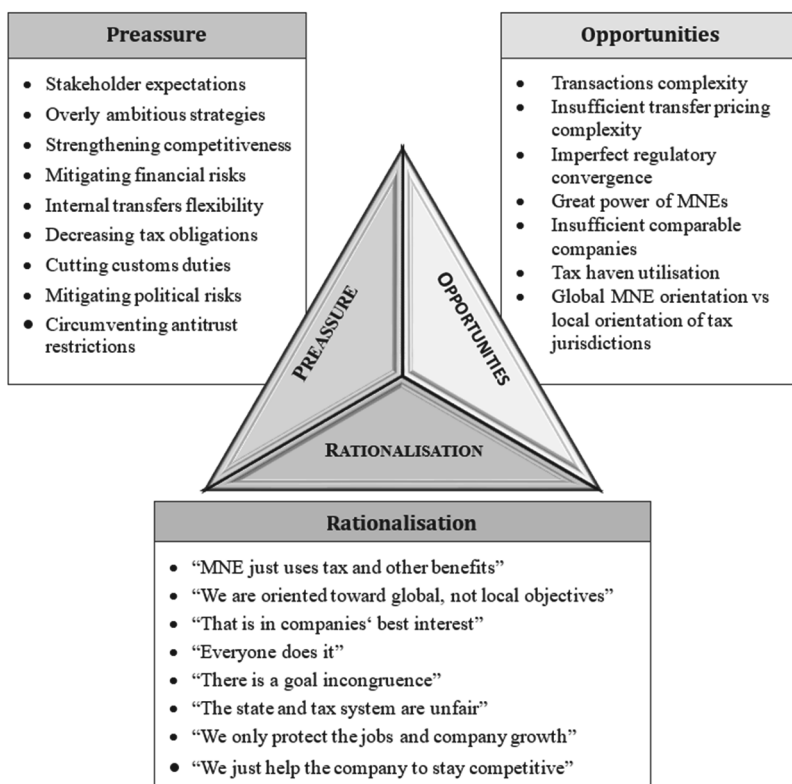
These risk factors can be further understood through the fraud triangle framework, which explains how pressure (reflecting motives for manipulation), opportunity (making fraud feasible) and rationalisation (eroding moral reservations about fraudulent activities) interact to increase the likelihood of fraudulent or opportunistic behaviour. Based on the framework, perpetrators commit fraud when they face problems that create pressure, identify opportunities or possibilities to carry it out and create a perception that such acts do not constitute as unacceptable behaviour (Cressey, 1950; Dorminey, Fleming, Kranacher, & Riley, 2012; Richardson, Taylor, & Wright, 2014). Given the numerous motives, and thus pressures to commit fraud, as well as various opportunities for its execution and the fact that rationalisation is an integral part of these processes, it stands to reason that the fraud triangle, as a general concept can adapt to each specific situation. In the context of potential transfer pricing manipulation, the fraud triangle manifests in specific ways within its core components, as illustrated in Figure 1.

There are various perspectives fraud triangle framework offers towards understanding transfer pricing fraud. For companies with profit and investment centres, the fraud triangle model is applied to identify critical points where opportunities for transfer pricing manipulation, primarily aimed at artificially inflating segment profitability, can be eliminated or at least minimised. For MNCs, the fraud triangle can be used to explore strategies that strengthen competitive advantages and maximise positive financial outcomes at the group level. Finally, from a regulator's perspective, considering the key components of the fraud triangle can help identify the motives and circumstances favourable to transfer pricing manipulation.

Transfer pricing related pressures typically derive from key stakeholders overly ambitious expectations. Such pressures may manifest as a need to strengthen

competitive position, for instance, by setting lower transfer prices in foreign markets to enhance competitiveness. Pressures to mitigate financial risk may drive decisions regarding intra-group financing under either favourable or unfavourable terms, depending on whether the aim is to retain profits domestically or shift them to a foreign tax jurisdiction. At times, transfer prices may be adjusted to mitigate risks associated with devaluation of local currency in the host country, thereby avoiding negative inflationary effects. Pressures will also manifest as demands to reduce tax costs, encouraging sales at lower transfer prices to jurisdictions with lower tax rates, and vice versa, thereby increasing group's overall profit in both scenarios. A similar scenario applies to the reduction of customs costs, where sales to related parties at lower transfer prices are preferred in order to mitigate impacts of customs duties (Lin & Chang, 2010). Aggressive investment in developing countries can bring political risks in the host countries, with transfer pricing once again emerging as a tool for mitigating such risks.

Figure 1: The Fraud Triangle - Manifestations of Pressure, Opportunity and Rationalisation in Transfer Pricing



Source: Authors

While pressures explain why the fraud occurs, opportunities can explain how it is conducted by highlighting the options available for carrying it out. In this context,

transfer pricing can serve as a useful instrument for manipulation execution. Opportunities for manipulation arise partly from the fact that transactions within large MNCs are often highly complex, particularly those involving digital goods, software licenses, know-how, etc. Perpetrators also exploit imperfect regulatory convergence, as regulations are not always adopted at the same pace which occasionally creates grey zones where transfer pricing manipulation becomes possible. The immense power of MNCs in such situations further facilitates the execution of these manipulations. It should be noted that the failure of national regulators to combat fraud in this area sometimes derives from the fact that MNC managers think globally, whereas regulators think and operate locally, overseeing individual entities that function within national borders. Ultimately, the effective management of risks associated with manipulation and fraud is significantly aided by a thorough understanding of vulnerabilities that can be exploited for fraudulent purposes, given that fraud cannot occur without such opportunities. Fraudulent opportunities are an indicator of a vulnerable environment where individuals and groups perceive an attractive potential for criminal gain, coupled with a low risk of detection and punishment (Ramamoorti, 2008). Only by recognising system vulnerabilities it is possible to implement robust policies and internal control systems to combat fraud (Oseifuah, 2025).

Ultimately, fraud also entails a justification for those who commit it. Perpetrators of fraud are humans who engage in manipulation, acting as agents on behalf of their organisations, so it is essential to try to understand their motives (Ramamoorti, 2008). Rationalisation in this sense has a behavioural dimension and explains how individuals justify their dishonest actions (Lederman, 2021). The reasons shown in Figure 1 are sufficiently clear, so we don't expore more detailed explanations in this paper. In fact, the fraud triangle emphasises traditional economic theory through the first two factors (pressure and opportunity), while the behavioral aspects of fraud become apparent through rationalisation.

Based on presented theoretical postulates, coupled with the empirical evidence, in this paper we focus on researching and developing a conceptual forensic framework to identify:

- To what extend identification of individual risk dimensions based on financial reporting information, can adequately assess total risk exposure to transfer pricing manipulation?
- Whether an integrated approach to assessing transfer pricing manipulation risk can alter a company's risk profile compared to an assessment based on individual risk dimensions?
- If utilisation of a multidimensional matrix can assist forensic accountants in assessing and nuancing company risk exposure to aggregated risk from transfer pricing manipulation?

More details are presented in the section below.

3. Forensic Transfer Pricing Manipulation Risk

Transactions within a group, characterised by common ownership, management and various forms of control, are subject to market mechanisms in the same way they would be if the entities operated as independent, standalone companies. In this context, transfer pricing serves as a mechanism for executing internal transfers between segments of a single company or between related parties within a group. Although complex and involving diverse arrangements among related parties, research on these relationships, in terms of risks of manipulation, typically focuses on product sales, use of intangible assets, services and intra-group financing (WCO, 2018).

In an effort to capitalise on the advantages of global market operations, companies can be expected to shift profits from high-tax jurisdictions to those with lower tax rates. Furthermore, it is realistic to expect that MNCs can utilise intra-group financing to enhance group's overall financial results. While tax related abuses are most obviously linked to manipulation and fraud, this does not rule out other motives and additional benefits of shifting profits to more tax favourable jurisdictions. Some of these activities may indeed be manipulative in nature. One must not overlook the risks associated with potential manipulations, such as avoiding or reducing customs duties, limiting competition, extracting liquid funds, misrepresenting the profitability of individual associated entities, manipulating management service charges, transferring intangible assets, internal debt relief, interest free lending, etc. Despite global efforts to prevent transfer pricing misuse, the empirical evidence clearly indicates that such practices persist.

The rapid development of digitalisation and artificial intelligence further complicates MNCs' operations oversight, thereby introducing new risks regarding the potential misuse of transfer pricing. The very fact that transfer pricing once involved physical goods, tangible assets and visible transactions, highlights the new challenges facing regulators in this area (Kelli, Brakye, Tamirepi, Timolijo, & Ulanja, 2026). It is clear that risks are escalating and their impact and interconnections are gaining complexity which requires integration of various risk factors aimed at identifying transfer pricing manipulations. Integrating these diverse risks, which may encompass digital intangible assets, varying regulatory requirements, business, financial and tax risks, complexities of international trade, reputational risks and the like, poses a significant challenge. Furthermore, the issue lies not only in individual risks but also in their scope regarding the dimensions they affect and their potential to amplify one another through interaction, which further complicates the problem.

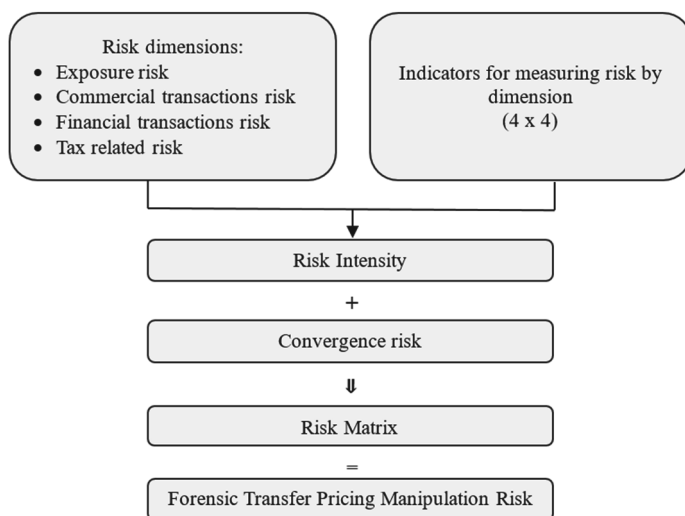
Given the complexity of potential transfer pricing manipulation, in this paper we provide a conceptual forensic framework that enables an assessment of various types of risks, including their potential interplay. In doing so, we proceed from the general definition of risk as a potential deviation from an expected outcome due to uncertainty. Since this deviation can be either positive or negative, risk can result in both opportunities and threats (ISO, 2022). As today, in terms of potential transfer

pricing manipulation, we cannot speak of exposure to a single type or dimension of risk, we proceed from the need for risk aggregation, providing a broader measure encompassing a range of distinct risks. In doing so, we rely theoretically on the general ISO definition of index aggregation, according to which “aggregate risk is the combination of different risks into a single risk to ensure a more comprehensive understanding of the total risk” (ISO, 2022).

3.1. Risk Intensity

Given the variety of issues associated with transfer pricing implementation process, the complexity of business transactions between related parties operating in different countries and the diversity of risks, the idea of risk aggregation becomes the core of measuring transfer pricing manipulation risk (Bjørnsen & Aven, 2019; Shivaani, Jain, & Yadav, 2019; Bao, Wana, Wua, & Li, 2019; Iossifov & Schmidt, 2021). In this paper, we adopt a conceptually broader approach to risk aggregation, aiming to identify a way to measure the risk associated with potential transfer pricing manipulation. Based on this, we operationalise a model involving two dimensional aggregation. First, we aggregate at the level of specific risk types that collectively reflect risk intensity, serving as a measure of the magnitude of a company's exposure to the risk of transfer pricing manipulation. Second, following the initial aggregation, a subsequent step combines this risk intensity with the number of distinct risk categories simultaneously present within the same business entity. We mark this as convergence risk. Figure 2 illustrates the model that integrates these concepts, moving towards the creation of a transfer pricing manipulation risk matrix.

To capture the key determinants of transfer pricing manipulation risk, we select four dimensions of risks inherent to transfer pricing: (1) risk of exposure to transactions with related parties (2) commercial transactions risk (3) financial transactions risk and (4) tax related risk.

Figure 2: Conceptual Model for Measuring Transfer Pricing Manipulation Risk

Source: Authors

The risk of exposure to transactions with related parties falls under the category of structural risks, as it relates to the structure and number of related parties and the international nature of the group. This risk partly covers transactions complexity which is inevitably greater when the group is more complex and transactions are more frequent. Significant transactions between related parties, international nature of such transactions or a concentration of transactions with a single related party or a small number of related parties can serve as warning signs of potential transfer pricing misuse.

Commercial transactions risk relates to the business activities of group members. While the mere existence of related parties doesn't automatically imply a high probability of manipulation, certain circumstances, such as complex transaction structures, transactions lacking a valid business purpose, heightened incentives to shift profits, certainly increase this risk. In this context, relationships between related parties offer greater opportunities to conceal potential transfer pricing manipulation. Factors such as presence of related parties in tax jurisdictions without economic activity, transactions involving intangible assets, management services and structuring of business transactions to achieve specific tax objectives serve as indicators that indicate the need for greater caution.

Financial transactions risk is associated with potential manipulation involving securing capital by intra-group financing. While such transactions can be legitimate, they may also be designed to shift financial results toward related parties in a way that generates tax or other benefits. Excessive interest expenses on intra-group financing, financing without or very little interest, fiduciary or solvency breaches,

inadequate debt capacity can indicate manipulation. In this context, changes in the capital structure can also serve as a warning sign.

Tax related risks refer to scenarios when companies account for incorrect amount of tax putting themselves into more favourable tax positions that may be incompatible with tax administrator's expectations. It may involve compliance risk resulting in fines and charges, reputational risk resulting from tax disputes and aggressive tax planning practices, transactional risk from specific transactions and pricing agreements or operational risk associated with internal processes that can have negative consequences on missed deadlines. Forensic red flags can be identified in company's profits or losses that are inconsistent with its business activities, global group's results inconsistent with tax burden over consecutive years, great disparities between effective and statutory tax rates, big volume of transactions without economic substance, interest rates appearing significantly higher compared to market, levels of indebtedness being disproportionately high compared to equity etc.

The advantage of presented risk intensity aggregation is in allowing us to view the overall risk while also seeing its composition, specifically, seeing whether it relates to one or multiple risk dimensions. This facilitates an understanding of the risk structure and identifies critical points for detecting potential transfer pricing manipulation.

3.2. Risk Intensity Indicators

In order to present fundamental ideas behind the concept of transfer pricing manipulation risk matrix, we decided to focus exclusively on financial indicators to measure the intensity of each individual risk dimension. This approach relies on information derived from financial statements, a positive aspect, given that all business and financial transactions are ultimately reflected in these statements and bear full impact of any potential transfer pricing manipulation. In a sense, this facilitates the acquisition of relevant information for model development. On the other hand, it must be acknowledged that certain non-financial indicators can also serve as significant signals of transfer pricing manipulation risk. Below, we present potential indicators, accompanied by brief descriptions on how they are interpreted in the context of possible transfer pricing manipulation.

Table 1: Risk Intensity Indicators

RD1 Risk of exposure to transactions with related parties		Interpretation
F11	Related parties revenue share	F11 growth indicates greater exposure to related parties transaction, increasing TP manipulation odds
F12	Related parties receivables share	F12 growth indicates greater exposure to related parties transaction, increasing TP manipulation odds
F13	Related parties liabilities share	F13 growth indicates greater exposure to related parties transaction, increasing TP manipulation odds
F14	Related parties intangibles growth	Greater F14 brings more challenges with defining ownership, contribution, real economic activities origin an comparability, leading to greater TP manipulation odds
RD2 Commercial transactions risk		Interpretation
F15	Sales revenue from related parties growth index	F15 growth increases TP manipulation odds
F16	Related parties receivables growth index	F16 growth increases TP manipulation odds
F17	Related parties collection period index	F17 growth indicates greater likelihood of overstating revenue and expenses, increasing TP manipulation odds
F18	Related parties gross operating margin index	When $F18 < 1$, gross operating margin is declining, decreasing TP manipulation odds

Table 1: Risk Intensity Indicators (continued)

RD3 Financial transactions risk		Indicator	RD3 Financial transactions risk	Interpretation
F19	Loan payable structure index		(Related parties loan payables _{it} / Total loan payables _{it}) / (Related parties loan payables _{it-1} / Total loan payables _{it-1})	F19 growth indicates greater intra-group financing, increasing TP manipulation odds
F110	Related parties financial expenses to financial revenue growth index		(Related parties financial expenses _{it} / Related parties financial revenue _{it}) / (Related parties financial expenses _{it-1} / Related parties financial revenue _{it-1})	When F110>1, expenses are growing more than revenues, in high-tax countries, increasing TP manipulation odds
F111	Thin capitalisation ratio		4 x average equity / Average debt ^{ff} from related parties	Greater F110 indicates greater internal debt financing, increasing TP manipulation odds
F112	Contracted to market interest rate		Contracted interest rate / Market interest rate	F112>1 indicates overstating financial results, while F112<1 indicates understating financial results, both leading to greater TP manipulation odds
RD4 Tax related risk		Indicator	RD4 Tax related risk	Interpretation
F113	Effective tax rate to statutory tax rate deviation		Effective tax rate / Statutory tax rate	With greater difference F113≠1, greater the need to examine the tax position, increasing TP manipulation odds
F114	Tax burden index		(Net Earnings _{it} / Earnings before tax _{it}) / (Net Earnings _{it-1} / Earnings before tax _{it-1})	Unusual F114 change may indicate a shift in tax behavior or tax structure, increasing TP manipulation odds
F115	Revenue to tax cost index		(Revenue _{it} / Revenue _{it-1}) / (Corporate income tax _{it} / Corporate income tax _{it-1})	When F115>1 corporate income tax doesn't keep pace with revenue growth, increasing TP manipulation odds
F116	Net deferred tax index		(Deferred tax assets _{it} / Deferred tax liabilities _{it}) / (Deferred tax assets _{it-1} / Deferred tax liabilities _{it-1})	F116>1 indicates growth of tax recoverables to payables suggesting greater tax planning activities, increasing TP manipulation odds

Source: Authors

A wide collection of diverse indicators further expands the scope of forensic investigations. We can now identify the specific determinants influencing each of the aforementioned risk dimensions. This enhances the ability to detect critical transactions that are potentially manipulated to achieve group objectives, regardless of the impact such behaviour might have on the performance of related parties.

3.3. Convergence Risk

Convergence in the context of transfer pricing can be viewed from different perspectives. It can be considered as a process of harmonisation between local and global legislative, tax systems and other regulatory compliance frameworks. In this context, convergent regulation reduces the risks of manipulative transfer pricing, thereby reducing tax risks from redirecting the tax burden towards more tax-friendly jurisdictions, as well as reducing risks of manipulative financial reporting that enables profitability shaping that deviates from economic reality.

In the context of this paper, convergence can be also viewed from the perspective of its long-term impact on the free market and global competition, driving the reduction of extreme profitability disparities within the same industries globally. However, this type of convergence can be undermined, particularly by MNCs, through transfer pricing manipulation, where related parties generate significant profitability differences that exceed deviations from the market acceptable profitability range. The magnitude of such manipulations is determined not only by the transfer prices themselves but also by the volume of transactions between related parties (Alex A.T. Rathke, Rezende, & Watrin, 2020).

The concept of risk convergence receives less attention in the literature, particularly in relation to transfer pricing, primarily due to intensity of risk being more varied, tangible and perhaps more compelling from a research perspective. However, in other areas, modern research involves an integrated risk management approach that implies that the analysis of individual risks must be supplemented by an examination of their simultaneous effects and interconnections, thereby demanding an understanding and management of the convergence existing between them (Cruz & Ayres da Silva, 2026).

The intensity of risk across various dimensions is undoubtedly the most important starting point for investigating transfer pricing manipulation risk. However, a comprehensive understanding of manipulation risk requires looking beyond individual risk dimensions and their aggregate score. One must also consider the extent to which these risk dimensions coexist within a company. For example, when transactions between related parties are complex, domestic company profitability is low, transactions involve related parties in low-tax jurisdictions and group-level margins are rising, then, in addition to the intensity of these individual risk factors, their simultaneous presence further increases company's exposure to the risk of transfer pricing manipulation. This creates a new, broader risk profile reflecting

heightened exposure to overall risk. Consequently, the probability of risk increases, as does the need to expand the scope of investigations into potential manipulation.

In this paper, we present convergence risk as a measure of the extent to which different risk dimensions manifest simultaneously across various business activities, thereby increasing the company's overall exposure to manipulation risk. It does not indicate the intensity of individual risks, but rather their horizontal spread. As such, convergence risk can also be understood as an increased probability of transfer pricing manipulation resulting from simultaneous or chronologically linked occurrence of multiple risk indicators from different dimensions.

Additionally, potential interactive effect of individual risks, which could, through synergy effects, amplify potential negative consequences should not be ignored. The possibility of an interactive effect between risk intensity and convergence risk implies that the interaction of these risks further increases overall exposure, requiring a broader analytical scope. This could potentially justify adding a third category to the existing risk classifications, risk interaction. However, given the scope and purpose of this paper, risk interaction will remain outside the scope of our analysis.

3.4. Risk Matrix

The risk matrix is a useful tool in which attributes of an unwanted event are being mapped to corresponding risk categories (Duijm, 2015). It is a convenient tool that can assist in the decision making process, sustainability analysis and risk assessment. There are two key questions that arise when creating any matrix. First, whether the matrix can satisfy user's interests and, second, whether the information in the matrix is reliable for decision making (Li, Bao, & Wu, 2019). In this sense, there is little doubt that quality warning signals are vital for forensic accountants, helping them target their investigations toward entities and activities most exposed to transfer pricing risks. This saves time in the investigation, reduces costs and increases the efficiency of forensic accountants. To the extent that such signals act as a restraint for potential perpetrators of fraud, they may also contribute to reducing some of its harmful consequences. This information is also important for group's management, that can get an idea of the size and consequences of playing transfer price games.

The existence of risk intensity and convergence risk is indisputable, as well as the possibility of constructing an index for each of these individual risks. However, we must point out that the nature of these risks is different, and therefore it wouldn't be easy to combine them in one measure, i.e., one index. Moreover, in doing so, we would probably lose more in differentiating potential situations in which companies can find themselves in terms of risk. We believe the forensic value would be greater if we preserved the two-dimensional visibility of these risk measures, which suggests a commitment to a risk matrix that would represent a combination of risk intensity and convergence risk. Operationalising the matrix requires defining a scale for both types of risk.

The creation of transfer pricing manipulation risk matrix should enable identification of risk nuances depending on the exposure to total risks. It is clear that risk intensity can vary between companies, depending on the different dimensions of risk to which companies are exposed. Combining different risk intensities by individual dimensions with different levels of convergence risk assumes their logical positioning in the matrix, which the risk scales make possible. In this context, when it comes to risk intensity, in Figure 3 we presented two scales, first a scale for individual indicators, and second a scale for realised risk scores that arise as a result of the effects of individual risk dimensions.

Figure 3: Risk Intensity

$Risk\ Intesity(RI) = \frac{\sum_{j=1}^{16} FI_j}{48}$	
<i>Risk Scale by Indicator</i>	
RI $\in \{0,1,2,3\}$	
0	Low risk signal
1	Moderate risk signal
2	Strong risk signal
3	Extremely high risk signal
<i>Risk Scale by Risk Dimension</i>	
0-12	Low risk signal
12-24	Moderate risk signal
24-36	Strong risk signal
36-48	Extremely high risk signal

Source: Authors

The first scale shows how strongly individual indicators send warning signals about exposure to transfer pricing manipulation risk. The second scale shows the strength of risk intensity at the aggregate level when all dimensions of risk that are subject to analysis, are included. It should be noted that the signal of low risk marked with zero should not be interpreted as a non-existent risk, but rather as an absence of materially significant risk manifestations. Given the relationship of these scales that represent same risks at different levels, and considering the intention to establish a conceptually clear model, we unified the scales in terms of risk interpretation, which allowed us to monitor risk intensity at four levels ranging from low to extremely strong exposure.

Risk intensity is measured as the contribution of the sum of risk intensity indicators' scores in all dimensions in the maximum score of all indicators (in our case, 4 dimensions x 4 indicators x 3 = 48). The index can be expressed and visualised across any interval within the range of 1% to 100%. The defined intervals therefore serve as categories for classifying index intensity, enabling a more meaningful and systematic interpretation of the results. When setting up a conceptual forensic model,

opting for equal intervals that represent thresholds for transitioning to higher risk categories is a good methodological choice due to them being easy to understand and methodologically neutral. However, empirical verification of the model will need to involve thresholds definition based on risk distribution in the sample, which can be done by using quartiles, percentiles, standard deviations or some other classification method.

As previously described, convergence risk indicates how widespread the risk is, in terms of how many different dimensions it is present in. In this sense, it is necessary to be careful with its interpretation for several reasons. First, there is a minimum risk intensity in risk dimension needed for convergence risk to be activated. The conceptual nature of the work allows us to assume that convergence risk will be activated when the aggregate risk intensity within one risk dimension equals 3. This means that there must be at least one warning signal of maximum intensity for one indicator or several lower intense indicators that in aggregate within one dimension activate the convergence. This approach is intended to prevent the activation of convergence risk at the level of one risk dimension due to some minimal deviation. At the same time, this approach gives forensic experts the freedom to evaluate the risk structure with the help of indicated points, depending on whether the risk is activated only under the influence of one or more indicators.

Determining the intervals in case of convergence risk has been simpler compared to risk intensity intervals. Convergence risk is different in its essence as it measures the prevalence of risk across dimensions in the sense that some (or all) dimension is active or not. As we adopted the four risk intervals for intensity, we also define four intervals for convergence risk, which we define as $1/4$ (or 25%, when only one dimension is activated), $2/4$ (50%), $3/4$ (75%) and $4/4$ (100%), that is, 25%, 50%, 75% or 100%, respectively, depending on whether one or more dimensions are activated. Hence, in this case too, we provide four levels of risk as can be seen in Figure 4.

Figure 4: Convergence Risk

Convergence - signal activation	
RD 1 $\geq$ 3	When the score within a single dimension is greater than or
RD 2 $\geq$ 3	equal to 3, the signal is activated and assigned a value of 1
RD 3 $\geq$ 3	When the score within a single dimension is less than 3, the
RD 4 $\geq$ 3	signal is not activated and assigned a value of 0
Convergence risk	
$\text{Convergence risk (CR)} = \frac{\text{No. risk dimension with score} \geq 3}{\text{Total no. dimension of risk}}$	
Risk Scale	
0-25%	Weak convergence
25-50%	Moderate convergence
50-75%	Strong convergence
75-100%	Total convergence

Source: Authors

In order to preserve the two-dimensional nature of risk, rather than pursuing a complex aggregation of intensity risk and convergence risk, which fundamentally differ in terms of their impact and magnitude, we opted for a conceptually more elegant solution in the form of a two-dimensional risk matrix, as presented in the Figure 5 below.

In order to set up the risk matrix, it was necessary to solve several important problems, partly in terms of risk classification depending on the effect of risk intensity and convergence risk, and partly also in terms of risk interpretation. In this sense, the presented matrix is based on several conditional rules.

Figure 5: Risk Matrix

Convergence risk Risk Intensity	Weak convergence	Moderate convergence	Strong convergence	Total convergence
Low risk signal	<i>Low risk</i>	<i>Low risk</i>	<i>Elevated risk</i>	<i>High risk</i>
Moderate risk signal	<i>Elevated risk</i>	<i>Elevated risk</i>	<i>High risk</i>	<i>High risk</i>
Strong risk signal	<i>High risk</i>	<i>High risk</i>	<i>High risk</i>	<i>Extremely high risk</i>
Extremely high risk signal	<i>Extremely high risk</i>	<i>Extremely high risk</i>	<i>Extremely high risk</i>	<i>Extremely high risk</i>

Source: Authors

First, the two key risk categories increase from the lowest to the highest level, with the thresholds set being methodologically neutral (Li, Bao, & Wu, 2019). Given the conceptual setting of the model, a descriptive presentation of the risk could not be avoided. Empirical verification of the model will require a slightly different approach.

Second, one of the main goals when creating the matrix was to appreciate the relevance of both risk categories, risk intensity and convergence risk. In doing so, it was assumed that different combinations of these risks can create asymmetric risk, i.e. risk that is disproportionately affected by risk intensity and convergence risk. For example, a low level of risk intensity cannot create extremely high risk in any variant, even if the convergence risk is at the highest level.

Third, the combination of risks in the matrix does not allow the highest level of risk intensity (extremely high risk) to be mitigated by the fact that the risk of convergence is present in only one, two or even all four dimensions. An extremely high level of risk intensity in any combination with convergence risk, will maintain the highest level of transfer pricing manipulation risk.

Fourth, for the same level of risk intensity, the convergence risk will have an amplifying effect on the overall risk. In this matrix structure, the effect will not be proportional. For example, at a low level of risk intensity, only strong convergence leads to increasing the risk to the level of elevated risk. Also, with moderate risk intensity, only strong convergence will increase the total risk to the level of high risk. Finally, high risk intensity level will lead to extremely high total risk only in combination with total convergence.

The risk matrix makes a moderate compromise between intensity and convergence, in a way that the convergence risk is not neglected, but it is given somewhat less importance. The potential impact of convergence risk on overall risk is such that it cannot be simply ignored. It is present in the matrix, but it does not have the power to turn low risk or even moderate risk intensity into extremely high total risk even at total convergence. In this sense, the matrix justifies the asymmetric effect of risk intensity and convergence risk. In fact, convergence risk in a certain combination with risk intensity can elevate overall risk, but in no combination can risk intensity be lowered due to convergence effects. While the usual requirement of same values along matrix diagonal is quite understandable, the asymmetry is found to be justified, bearing in mind the different nature of the risk we are analysing (Li, Bao, & Wu, 2019).

Presented matrix largely fulfils Cox's criteria by providing clearly recognisable extreme values (highs and lows), by not supporting sudden changes from low to highest priority and by the fact that similar risk exposures belong to similar categories (Cox, Jr, 2008). Potential deviation can be identified in the fact that the highest level of risk intensity cannot be reduced by any level of convergence risk, which is understandable if we keep in mind the nature of these risks.

Certainly, the risks of using these types of matrices must not be ignored. They are especially present if the dimensions of risk are not well defined, if the thresholds are incorrectly set, if biases in matrix creation is pronounced or if a score is treated as an absolutely accurate measure of risk. Finally, it should be borne in mind that risk matrices must correspond to specific circumstances, in this case transfer pricing manipulation, so that the specificities of some matrix are acceptable and even desirable, unless they deviate from some generally accepted criteria that we mentioned previously.

Conclusion

While globalisation of business brings MNCs great opportunities to realise numerous advantages in international trade, it also carries many risks that can lead to serious consequences. Transfer pricing as a mechanism for valuing transfers between related parties is sometimes misused to achieve tax and other benefits at the group level by shifting profits to countries with more favourable tax rates. At the same time, it should be noted that the existence of related parties and transfer pricing doesn't automatically mean a high risk exposure to manipulation. However, a combination of different circumstances, such as existence of complex group structures, significant transactions outside of core business, transfers of digital assets, business operations in tax-attractive industries while value is created in other countries, etc. can lead to a significant increase in risk of manipulation.

Having a comprehensive overview of such risks is essential as transfer pricing can be found at the intersection of taxation regulation, financial reporting, corporate governance and competitive strategies. Failure to holistically monitor transfer pricing risk exposure creates severe vulnerabilities across companies, markets and even countries. Providing analytical instruments that enable forensic specialists to detect true economic value, identify transfer pricing manipulation risk and detect pricing deviations before fraudulent behaviour occurs becomes crucial.

The provided model for identifying risk of transfer price manipulation, which includes risk intensity and convergence risk, enables a better understanding of the need to design instruments that contribute to preventing transfer pricing misuse. The model presented in the paper offers one, we would say useful conceptual tool for further analysis. Its empirical confirmation would be a step forward in improving the detection and prevention of transfer pricing manipulation.

Finally, we must point out the limitations that are present in the work. Since our goal was to create a conceptual forensic framework for determining exposure to the risk of manipulation, we did not have space for empirical verification of some key points. The created conceptual approach is understandable, consistent and methodologically neutral, but only empirical verification of this model can identify at what level the thresholds should be set, how many dimensions of risk intensity should be taken into consideration, how individual indicators should be evaluated and how to classify risk levels in the matrix. These are also directions for further research into risk management in the area of transfer price manipulation.

References

- Alex A.T. Rathke, A. A., Rezende, A. J., & Watrin, C. (2020). The impact of countries' transfer pricing rules on profit shifting. *Journal of Applied Accounting Research*, 1-28. Retrieved from <https://www.emerald.com/insight/0967-5426.htm>
- Alghamdi, F., Eulaiwi, B., Hussein, S., Duong, L., & Taylor, G. (2024). Non-arm's-length transactions, offshore financial centres, transfer pricing agreements and corporate cash holdings: Evidence from U.S. multinational corporations. *International Review of Financial Analysis*, 96(1). <https://doi.org/10.1016/j.irfa.2024.103589>
- Arnold, M., Elsinger, F., & Rankin, F. (2021). The Unintended Consequences of Headquarters' Involvement in Decentralized Transfer Price Negotiations: Experimental Evidence. *Management Science*, 67(12), 7912-7931. <https://doi.org/10.13140/RG.2.2.33786.29125>
- Bao, C., Wana, J., Wua, D., & Li, J. (2019). Aggregating risk matrices under a normative framework. *Journal of Risk Research*, 1-18. Retrieved from <https://doi.org/10.1080/13669877.2019.1588912>
- Bjørnsen, K., & Aven, T. (2019). Risk aggregation: What does it really mean? *Reliability Engineering and System Safety* (2019). Retrieved from doi: <https://doi.org/10.1016/j.ress.2019.106524>
- Buettner, T., & Wamser, G. (2013). Internal Debt and Multinational Profit Shifting: Empirical Evidence from Firm-Level Data. *National Tax Journal*, 66(1), 63-96. <https://doi.org/10.17310/ntj.2013.1.03>
- Buettner, T., Overesch, M., & Wamser, G. (2018). Anti Profit-Shifting Rules and Foreign Direct Investment. *International Tax and Public Finance*, 25(3), 553-580.
- Contreras, L., Hugger, F., & Zawisza, T. (2026). MNE Responses to the Global Minimum Tax. *OECD Taxation Working Papers*, 77. <https://doi.org/10.1787/0bc9aac9-en>
- Cox, Jr, L. A. (2008). What's Wrong with Risk Matrices? *Risk Analysis*, 28(2), 497-512. <https://doi.org/10.1111/j.1539-6924.2008.01030.x>
- Cressey, D. R. (1950). The criminal violation of financial trust. *American Sociological Review*, 15, 738, 19, 738-743.
- Cristea, A. D., & Nguyen, D. X. (2016). Transfer Pricing by Multinational Firms: New Evidence from Foreign Firm Ownerships. *American Economic Journal: Economic Policy*, *American Economic Association*, 8(3), 170-202. <https://doi.org/10.1257/pol.20130407>
- Crivelli, E., De Mooij, R., & Keen, M. (2015). *Base Erosion, Profit Shifting and Developing Countries*. IMF Working Paper 15/118. <https://doi.org/10.5089/9781513563831.001>

- Cruz, C. O., & Ayres da Silva, A. L. (2026). Risk convergence in mining: Climate change as a multiplier of operational, geotechnical, and reliability threats. *The Extractive Industries and Society*, 28, 1-11.
- Davies, R., Martin, J., Parenti, M., & Toubal, F. (2018). Knocking on Tax Haven's Door: Multinational Firms and Transfer Pricing. *The Review of Economics and Statistics*, 100(1), 120-134. https://doi.org/10.1162/REST_a_00673
- De Mooij, R., & Liu, L. (2020). At A Cost: the Real Effects of Transfer Pricing Regulations. *IMF Economic Review*, 68(1), 268-306.
- De Simone, L., Klassen, K., & Seidman, J. (2017). Unprofitable Affiliates and Income Shifting Behavior. *The Accounting Review*, 92(3), 113–136. <http://dx.doi.org/10.2308/accr-51555>
- DiMaggio, P., & Powell, W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in. *American Sociological Review*, 48(2), 147-160. Retrieved from <http://www.jstor.org/stable/2095101>
- Dorminey, J., Fleming, S. A., Kranacher, M.-J., & Riley, R. A. (2012). The Evolution of Fraud Theory. *Issues in Accounting Education*, 29(32), 571.
- Duijm, N. J. (2015). Recommendations on the use and design of risk matrices. *Safety Science*, 76, 21-31. <http://dx.doi.org/10.1016/j.ssci.2015.02.014>
- Eccles, R. (1985). *The Transfer Pricing Problem: A Theory for Practice*. New York: Lexington Books.
- Elexa, L., Ištók, M., & Šlampaiová, L. (2022). Do Links to Tax Havens Affect Companies' Financial Performance? The Case of Slovakia. *E&M Economics and Management*, 25(1), 60-76. [10.15240/tul/001/2022-1-004](https://doi.org/10.15240/tul/001/2022-1-004)
- Eukeria, W., & Mpofo, F. Y. (2024). Manipulation of Transfer Pricing Rules by Multinational Enterprises in Developing Countries: The Challenges and Solutions. *Journal of Tax Reform*, 10(1), 181-207. <https://doi.org/10.15826/jtr.2024.10.1.164>
- Fuest, C., Hebous, S., & Riedel, N. (2011). International Debt Shifting and Multinational Firms in Developing Countries. *Economics Letters*, 113(2), 135-38. <https://doi.org/10.1016/j.econlet.2011.06.012>
- Fuest, C., Hugger, F., & Neumeier, F. (2022). Corporate profit shifting and the role of tax havens: Evidence from German country-by-country reporting data. *Journal of Economic Behavior & Organization*, 194(C), 454-477. <https://doi.org/10.1016/j.jebo.2021.11.016>
- Garcia-Bernardo, J., & Janský, P. (2024). Profit shifting of multinational corporations worldwide. *World Development*, 177. <https://doi.org/10.1016/j.worlddev.2023.106527>
- Gauß, P., Kortenhaus, M., Riedel, N., & Simmler, M. (2024). Leveling the playing field: Constraints on multinational profit shifting and the performance of national

firms. *Journal of Public Economics*, 234.
<https://doi.org/10.1016/j.jpubeco.2024.105116>

Huizinga, H., Laeven, L., & Nicodème, G. (2008). Capital Structure and International Debt Shifting. *Journal of Financial Economics*, 88(1), 80-118.
<https://doi.org/10.1016/j.jfineco.2007.05.006>

Iossifov, P., & Schmidt, T. D. (2021). *Cyclical Patterns of Systemic Risk Metrics: Cross-Country Analysis*. International Monetary Fund.
<https://doi.org/10.5089/9781513568652.001.A001>

ISO. (2022). *Risk management—Vocabulary, ISO 31073*. International Organization for Standardization. Preuzeto sa <https://www.iso.org/standard/79637.html>

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](https://doi.org/10.1016/0304-405X(76)90026-X)

Johannessen, N., Tørsløv, T., & Wier, L. (2020). Are Less Developed Countries More Exposed to Multinational Tax Avoidance? Method and Evidence from Micro-Data. *The World Bank Economic Review*, 34(3), 790–809.
<https://doi.org/10.1093/wber/lhz002>

Kalra, A., & Afzal, M. (2023). Transfer pricing practices in multinational corporations and their effects on developing countries' tax revenue: a systematic literature review. *International Trade, Politics and*, 7(3), 172-190.
<https://doi.org/10.1108/ITPD-04-2023-0011>

Kelli, A., Brakye, K., Tamirepi, L., Timolijo, B., & Ulanja, M. (2026). Designing policy frameworks for addressing transfer pricing risks in digital and data-driven business models. *Journal of Accounting and Taxation*, 18(2), 55-72. Retrieved from <https://doi.org/10.5897/JAT2025.0673>

Klassen, K., Lisowsky, P., & Mescall, D. (2017). Transfer Pricing: Strategies, Practices, and Tax Minimization. *Contemporary Accounting Research*, 34(1), 455-493. <http://dx.doi.org/10.2139/ssrn.2216870>

Laudage Teles, S., Riedel, N., Strohmaier, K., Strohmaier, K., Voget, J., & Wickel, S. (2026). Transfer price documentation rules and multinational firm behavior - evidence from France. *Journal of Public Economics*, 257.
<https://doi.org/10.1016/j.jpubeco.2026.105565>

Lederman, L. (2021). The fraud triangle and tax evasion. *Iowa Law Review*, 1153–1207. Retrieved from <https://www.repository.law.indiana.edu/facpub/2968>

Li, J., Bao, C., & Wu, D. (2019). How to Design Rating Schemes of Risk Matrices: A Sequential Updating Approach. *Risk Analysis*, 1-19.
<https://doi.org/10.1111/risa.12810>

Lin, C.-W., & Chang, H.-C. (2010). Motives of transfer pricing strategies – systemic analysis. *Industrial Management & Data Systems*, 110(8), 1215-1233. <https://doi.org/10.1108/02635571011077843>

Liu, L., Schmidt-Eisenlohr, T., & Guo, D. (2020). International transfer pricing and tax avoidance: Evidence from the linked tax-trade statistics in the United Kingdom. *The Review of Economics and Statistics*, 102(4), 766-778. https://doi.org/10.1162/rest_a_00871

Moyo, N. M., & Liu, Y. (2022). Transfer Pricing As a Tool for Tax Avoidance by Multinational Enterprises in Developing Countries. *Journal of Business and Management*, 24(10), 16-32.

OECD. (2022). *OECD Transfer Pricing Guidelines for Multinational Enterprises and Tax Administrations 2022*. Paris: OECD Publishing. <https://doi.org/10.1787/0e655865-en>

OECD. (2026). Corporate income taxation and business dynamism. *OECD Tax Policy Briefs*, 2026/01. <https://doi.org/10.1787/59a40a50-en>

Oseifuah, E. K. (2025). Application of the fraud triangle model in mitigating tax evasion in developing countries: a conceptual framework. *Business Ecosystem & Strategy*, 7(2), 297-306. <https://www.bussecon.com/ojs/index.php/ijbes>

Ramamoorti, S. (2008). The Psychology and Sociology of Fraud: Integrating the Behavioral Sciences Component Into Fraud and Forensic Accounting Curricula. *Issues in Accounting Education*, 23(4), 521-533.

Richardson, G., Taylor, G., & Wright, C. (2014). Corporate profiling of tax-malfeasance: A theoretical and empirical assessment of tax-audited Australian firms. *eJournal of Tax Research*, 12(2), 359-382.

Shivaani, M. V., Jain, P. K., & Yadav, S. S. (2019). In N. F. Analysis, *Understanding Corporate Risk. India Studies in Business and Economics*. Singapore: Springer. Retrieved from https://doi.org/10.1007/978-981-13-8141-6_3

Tørsløv, T., Wier, L., & Zucman, G. (2023). The Missing profits of nations. *The Review of Economic Studies*, 90(3), 1499-1534. <https://doi.org/10.1093/restud/rdac049>

Viertola, M. (2023). *Profit Shifting of Multinational Enterprises: Evidence from the Nordics, Working Paper 18*. Helsinki: Finnish Center of Excellence in Tax Systems Research. Retrieved from <https://verotutkimus.fi/verotutkimus/wp-content/uploads/2023/12/FIT-WP-18-Profit-Shifting-of-Multinational-Enterprises.pdf>

WCO. (2018). *WCO Guide to Customs Valuation and Transfer Pricing*. World Customs Organization.

Wier, L. (2020). Tax-motivated transfer mispricing in South Africa: Direct evidence using transaction data. *Journal of Public Economics*, 184. <https://doi.org/10.1016/j.jpubeco.2020.104153>

Ivan Dionisijev
Ss. Cyril and Methodius University in
Skopje, Faculty of Economics – Skopje
ORCID: 0000-0001-5986-3343

Zorica Bozhinovska Lazarevska
Ss. Cyril and Methodius University in
Skopje, Faculty of Economics – Skopje
ORCID: 0000-0002-9074-0365

Monika Mitrevska
Ss. Cyril and Methodius University in
Skopje, Faculty of Economics – Skopje

PROFESSIONAL SKEPTICISM AND THE IDENTIFICATION OF IRREGULARITIES IN PUBLIC SECTOR AUDITS: EVIDENCE FROM STATE AUDITORS

UDK 657.6:351.9

DOI <https://doi.org/10.54318/czf.2026.id.004>

Abstract

Professional skepticism is a fundamental element of auditing, supporting the critical evaluation of audit evidence, sound professional judgment, and the identification of circumstances requiring further scrutiny. While professional skepticism has been extensively examined in private-sector auditing, considerably less empirical attention has been devoted to its role within Supreme Audit Institutions (SAIs). This study examines professional skepticism among state auditors and its relationship with their perceived ability to identify significant irregularities in public sector audits. It further considers organizational support, professional training and experience, and perceived organizational and institutional pressures as contextual factors surrounding skeptical audit behavior. The study employed a quantitative, cross-sectional research design based on a structured questionnaire completed by 35 state auditors employed by the State Audit Office of the Republic of North Macedonia. The results indicate a high reported level of professional skepticism and reveal a strong positive association between professional skepticism and auditors' perceived ability to identify significant irregularities. All individual skeptical behaviors were positively and significantly associated with irregularity identification, with the strongest relationships observed for avoiding premature conclusions, actively searching for contradictions and inconsistencies, and considering alternative explanations. The findings extend behavioral auditing research to the comparatively underexplored SAI environment and highlight professional skepticism as an important behavioral attribute associated with irregularity-sensitive auditing and public sector accountability.

Key words: professional skepticism; public sector auditing; Supreme Audit Institutions; irregularity identification; auditor judgment; forensic accounting

JEL Classification: M42; H83; K42

Introduction

Public sector auditing plays an essential role in safeguarding public resources and strengthening accountability, transparency, and trust in public institutions. Supreme Audit Institutions (SAIs), as independent external audit bodies, contribute to this process by examining the use of public resources and providing legislatures, public institutions, and citizens with independent assurance on government activities. At the same time, the boundaries of public sector auditing have expanded beyond traditional financial and compliance auditing toward performance auditing, sustainability-related issues, and increasingly digitalized audit environments (Hay & Cordery, 2018; Grossi et al., 2023; Volodina & Grossi, 2025). These developments increase both the complexity of audit engagements and the importance of auditors' ability to critically assess diverse forms of evidence and exercise sound professional judgment.

Within this environment, professional skepticism represents a fundamental behavioral attribute of audit work. It requires auditors to maintain a questioning mind, remain alert to conditions that may indicate misstatement or irregularity, critically evaluate audit evidence, and avoid premature conclusions when the available information is insufficient or contradictory (Nelson, 2009; Hurtt, 2010). Such behavior becomes particularly important when auditors encounter unusual transactions, inconsistent explanations, contradictory evidence, or other warning signs that may require additional examination. Professional skepticism can therefore contribute not only to the quality of professional judgment but also to auditors' ability to recognize significant irregularities and circumstances requiring further audit scrutiny.

Previous empirical research provides evidence of a positive relationship between professional skepticism and fraud- or irregularity-related audit outcomes. Studies involving internal and government auditors indicate that skeptical behavior can strengthen auditors' ability to recognize problematic circumstances and fraud-related indicators (Agustina et al., 2021; Sari & Andrian, 2023; Rininda, 2024). However, much of the empirical literature on professional skepticism has focused on private-sector external auditors or government internal auditors, while evidence from auditors employed by SAIs remains comparatively limited. This represents an important gap because state auditors operate under a distinct institutional mandate characterized by external public-sector oversight, broad accountability responsibilities, diverse stakeholders, and potentially sensitive audit environments.

North Macedonia provides a particularly relevant setting in which to examine this issue. The State Audit Office (SAO) has undergone substantial institutional development since becoming operational in 1999, while its mandate and professional capacity have expanded alongside broader developments in public sector auditing (Dionisijev et al., 2022; Dionisijev & Bozhinovska Lazarevska, 2024). Although previous research in the Macedonian context has examined professional skepticism among internal auditors in relation to fraud indicators (Bozhinovska Lazarevska et

al., 2017), considerably less is known about the skeptical behavior of state auditors employed by the country's SAI and, more specifically, whether professional skepticism is associated with their perceived ability to identify significant irregularities.

Against this background, the primary objective of this study is to examine the relationship between professional skepticism and state auditors' perceived ability to identify significant irregularities during public sector audits. The study additionally considers professional training and experience, organizational support, and perceived organizational and institutional pressures as contextual factors surrounding skeptical audit behavior. To address these objectives, empirical evidence is obtained from state auditors employed by the Macedonian SAO and analyzed using descriptive statistics, reliability analysis and correlation analysis.

The study contributes to the existing literature in several respects. First, it extends research on professional skepticism to the comparatively underexplored setting of a SAI. Second, it shifts attention from professional skepticism as a general audit attribute toward the specific skeptical behaviors associated with the identification of significant irregularities, including questioning information, searching for contradictions, considering alternative explanations, and withholding conclusions until sufficient evidence has been obtained. Third, by providing evidence from North Macedonia, the study contributes to the limited behavioral auditing literature from smaller and less frequently examined public-sector environments. Finally, it establishes a relevant connection between public sector auditing and forensic accounting by examining behavioral mechanisms that may facilitate the recognition of irregularities and warning signs, while maintaining the important distinction between identifying such indicators and formally establishing or investigating fraud.

1. Literature review

1.1. Public sector auditing and SAIs

Public sector auditing is an important mechanism for strengthening accountability, transparency, and control over public resources. Its value extends beyond verification of financial information because public auditors may also assess compliance, economy, efficiency, and effectiveness. Hay and Cordery (2018) explain public audit through several complementary perspectives, particularly agency and management control. In the public sector, multiple principal-agent relationships, from citizens and legislatures to governments and public organizations, create information asymmetries and a need for independent assurance. SAIs address this need by providing external oversight of government activities and the use of public funds, thereby supporting accountability and good governance (Hay & Cordery, 2018; Bonollo, 2019; Dionisijev & Bozhinovska Lazarevska, 2024).

The scope of public sector auditing has expanded considerably. SAIs have moved from a predominantly financial and compliance-oriented role toward performance

auditing and broader assessments of public policies and programs. Grossi et al. (2023) describe changing boundaries of public sector auditing associated with performance, sustainability, and digitalized auditing, while Cordery and Hay (2025) show that New Public Management and subsequent reforms have altered both the context and activities of public auditors. These developments increase the complexity of audit assignments and require auditors to evaluate more diverse evidence and risks. Digital transformation adds another dimension: Volodina and Grossi (2025) show that technological change may support public value and accountability, but it also creates concerns and new demands for auditors' competencies and judgment. More broadly, Samiolo et al. (2024) demonstrate that technological transformation also raises important questions about the nature and development of auditor judgment, as automation may alter opportunities for deliberation, sensemaking, and professional learning.

This institutional evolution is relevant to North Macedonia. The SAO began operating as the country's SAI in 1999 and subsequently became a member of INTOSAI and EUROSAI. Its mandate includes regularity audits and, since 2005, performance audits. Previous research documents the SAO's organizational development and emphasizes the importance of professional capacity, training, international experience, institutional support, and effective follow-up of audit recommendations (Dionisijev et al., 2022). More recent work also highlights the role of SAIs in sustainability-related auditing and the Macedonian SAO's engagement with the Sustainable Development Goals (Dionisijev & Bozhinovska Lazarevska, 2024). As public audit becomes broader and more complex, understanding the behavioral qualities of individual state auditors becomes increasingly important.

1.2. Professional skepticism

Professional skepticism is a central concept in auditing because it guides how auditors approach information and evaluate evidence. Nelson (2009) conceptualizes skepticism through skeptical judgments and actions that reflect a heightened assessment of the possibility that an assertion may be incorrect. His model shows that skeptical judgment is influenced by evidential inputs, auditor knowledge, individual traits, incentives, and the audit environment. Professional skepticism therefore does not imply automatic distrust; rather, it requires auditors to remain alert, question available information, and demand persuasive evidence before accepting an audit conclusion.

At the same time, the literature recognizes different perspectives on how skepticism should be exercised in practice. Glover and Prawitt (2014) conceptualize professional skepticism as a continuum whose appropriate application may vary with audit risk, while Quadackers et al. (2014) distinguish between neutrality and presumptive doubt and find the latter to be more predictive of skeptical judgments and decisions, particularly in higher-risk settings.

Hurt (2010) provides a complementary trait-based perspective and identifies six characteristics of professional skepticism: a questioning mind, suspension of judgment, search for knowledge, interpersonal understanding, self-esteem, and autonomy. The first three characteristics relate specifically to how auditors examine evidence, reflecting a willingness to seek and fully evaluate sufficient evidence before reaching a judgment. Subsequent literature also distinguishes between relatively stable trait skepticism and context-dependent state skepticism, indicating that skeptical behavior can be affected by both personal characteristics and situational conditions (Hurt et al., 2013). The relevance of these characteristics has also been examined in the public sector. Ghani et al. (2022), applying Hurt's framework to public sector accountants in Malaysia, found that a questioning mind, suspension of judgment, and search for knowledge were particularly relevant dimensions of professional skepticism.

Recent research increasingly treats professional skepticism as a behavioral and context-sensitive process rather than only a fixed personal characteristic. Auditors interpret cues, interact with colleagues and audited entities, and revise their understanding as new evidence emerges. Empirical studies generally support the importance of skepticism for audit judgment. Shirowzhan and Fakhari (2024) report a positive association between professional skepticism and the quality of auditors' judgment and decision-making, while Budiadnyani and Dewi (2024) similarly find that professional skepticism improves audit judgment. At the same time, evidence that the relationship varies across contexts suggests that skepticism should be examined together with experience, organizational conditions, and pressures rather than in isolation.

Most empirical skepticism research has been conducted in private-sector or external audit settings. Evidence from government and SAI environments is more limited, even though state auditors operate in settings characterized by public accountability, political sensitivity, complex mandates, and diverse stakeholders. This makes the public sector an important context for examining whether auditors maintain a questioning approach, seek additional evidence, remain alert to irregularities, and revise conclusions when new information becomes available.

1.3. Professional skepticism and the identification of irregularities

Professional skepticism becomes particularly important when auditors encounter information that may indicate irregularities, material misstatements, or potentially fraudulent activity. A questioning mindset encourages auditors to critically evaluate the reliability of information, seek corroborating evidence, investigate inconsistencies, and consider alternative explanations before reaching a conclusion. Professional skepticism can therefore be viewed as an important behavioral link between the evaluation of audit evidence and the recognition of circumstances that require additional audit attention.

Prior empirical research supports the relevance of professional skepticism for the identification of fraud and related risk indicators. Agustina et al. (2021), examining internal auditors in the public sector, investigated professional skepticism in the relationship between auditor competence, independence, experience, time pressure, and fraud detection. Their findings underline the importance of skeptical behavior in translating auditors' professional capabilities into fraud-detection performance. Similarly, Sari and Andrian (2023) found a significant positive relationship between professional skepticism and fraud detection, while auditor experience did not have a statistically significant effect. These findings suggest that experience alone may not be sufficient and that the way auditors critically approach and evaluate evidence represents an important element of their ability to recognize problematic circumstances.

Evidence from government audit environments provides further support for this relationship. Rininda (2024), based on a sample of 163 government internal auditors, found that both professional skepticism and audit experience positively affected auditors' ability to detect fraud. Using a qualitative approach, Butar Butar (2018) examined government forensic auditors and showed how characteristics of professional skepticism, including a questioning mind, suspension of judgment, search for knowledge, interpersonal understanding, self-confidence, and self-determination, were applied in fraud-detection activities. Taken together, these studies indicate that skeptical behavior is particularly relevant in audit situations involving uncertainty, contradictory evidence, unusual circumstances, and potential warning signs.

Evidence from North Macedonia provides an important national antecedent. Bozhinovska Lazarevska et al. (2017) examined professional skepticism among internal auditors in the context of fraud identification, focusing specifically on whether individual levels of skepticism influence auditors' behavior toward fraud indicators during audit engagements. Their research represents an important early attempt to examine the behavioral dimension of professional skepticism in the Macedonian environment. However, its focus was on internal auditors rather than state auditors employed by the country's SAI. This distinction is important because state auditors operate under a different institutional mandate, conduct external public sector audits, and exercise professional judgment in an environment characterized by broader public accountability responsibilities.

At the same time, identifying irregularities during a public sector audit should not be equated with formally establishing or investigating fraud. Public sector auditors may encounter inconsistencies, unusual transactions, material misstatements, insufficient explanations, or other circumstances that require additional evidence and further audit procedures without determining that fraudulent conduct has occurred. In the context of the present study, the relevance of professional skepticism therefore lies primarily in its potential to strengthen auditors' sensitivity to irregularities and risk indicators and their willingness to subject questionable evidence to further scrutiny.

1.4. Determinants of professional skepticism and irregularity identification

Professional skepticism and the ability to recognize irregularities may be shaped by a combination of individual, professional, organizational, and situational factors. Auditor experience represents one such factor. Repeated exposure to different audit engagements can develop auditors' knowledge of common risk patterns, inconsistencies, unusual transactions, and explanations provided by audited entities. Nelson (2009) incorporates auditor knowledge and experience into the process through which skeptical judgments and actions are formed. Empirical evidence, however, is not entirely consistent. Rininda (2024) reports a positive effect of experience on fraud-detection ability among government internal auditors, while Sumartono et al. (2022), studying government external auditors within an SAI environment, found that both auditor experience and professional skepticism were associated with audit judgment. In contrast, Sari and Andrian (2023) find no significant independent effect of auditor experience. Such mixed findings indicate that accumulated experience should not automatically be equated with stronger irregularity-identification ability.

Professional training and continuous development may complement experience by strengthening technical competence and auditors' ability to recognize emerging risks and unusual evidence. This is particularly relevant in public sector auditing, where auditors increasingly deal with complex regulatory requirements, performance information, sustainability issues, and digitalized audit environments. Previous research on the Macedonian SAO has similarly emphasized professional capacity, staff training, and international cooperation as important elements in the institutional development of the SAI (Dionisijev et al., 2022). Training may therefore support both individual audit competence and the broader institutional capacity required for high-quality professional judgment.

Organizational support represents another important contextual factor. Skeptical judgment does not necessarily translate automatically into skeptical action. Auditors may recognize that explanations are incomplete or that additional evidence is necessary, but their willingness and ability to pursue further procedures may depend on the organizational environment. Nelson (2009) emphasizes that incentives and environmental conditions influence both skeptical judgment and skeptical action. An environment that encourages critical thinking, supports additional questioning and evidence gathering, and protects independent professional judgment may consequently facilitate the practical application of professional skepticism.

Conversely, time and organizational pressures may constrain skeptical audit behavior. Short deadlines and limited resources can reduce the opportunity to pursue additional evidence or investigate inconsistencies in greater depth. Consistent with this concern, Sayed Hussin et al. (2017) found that time budget pressure negatively affected auditors' assessment of misstatement risk and weakened the positive relationship between professional skepticism and risk assessment, whereas audit experience strengthened that relationship. Agustina et al. (2021) explicitly consider

audit time pressure together with professional skepticism in the context of public sector fraud detection, while the broader skepticism literature recognizes that situational factors can affect whether skeptical judgments result in corresponding audit actions. Organizational hierarchy, external expectations, and politically or institutionally sensitive circumstances may also influence the audit environment, particularly in the public sector.

The existing evidence therefore suggests that professional skepticism should not be examined in isolation. Experience, training, organizational support, and perceived pressures provide a broader professional and institutional context within which skeptical behavior is exercised. At the same time, inconsistent findings concerning their independent effects justify examining these factors empirically rather than assuming that they necessarily improve or weaken auditors' ability to identify irregularities.

1.5. Professional skepticism, irregularity identification and public sector accountability

The relationship between professional skepticism and irregularity identification has particular significance in public sector auditing. SAIs operate within complex accountability relationships involving governments, legislatures, public organizations, and citizens, and the credibility of their findings depends on the quality of the evidence and professional judgments underlying audit conclusions (Hay & Cordery, 2018). As the boundaries of public sector auditing expand beyond traditional financial and compliance auditing toward performance, sustainability, and increasingly digitalized audit environments, the ability to critically assess diverse forms of evidence becomes even more important (Grossi et al., 2023; Volodina & Grossi, 2025).

For SAIs, professional skepticism may therefore contribute to the reliability of audit findings and broader public accountability by strengthening the evidence-evaluation process through which potentially significant irregularities are recognized.

The present study addresses this gap by examining professional skepticism among state auditors in North Macedonia and its association with their perceived ability to identify significant irregularities during public sector audits. It further considers professional training and experience, organizational support, and perceived pressures as contextual factors surrounding skeptical audit behavior. In doing so, the study extends the behavioral auditing literature to a comparatively underexplored SAI environment and provides a bridge between research on professional skepticism, public sector auditing, and the identification of irregularities relevant to forensic accounting.

2. Research methodology

This study employed a quantitative, cross-sectional research design to examine professional skepticism among public sector auditors and its relationship with the identification of irregularities in public sector audits. Primary data were collected through a structured questionnaire distributed to state auditors employed by the Macedonian SAO. A total of 35 valid responses were obtained and included in the empirical analysis.

The questionnaire consisted of demographic and professional questions and 30 statements (Annex 1) grouped into five dimensions:

- 1) Professional Skepticism (PS), measured by ten items;
- 2) Auditor Judgment (AJ), measured by four items;
- 3) Organizational Support (OS), measured by five items;
- 4) Perceived Organizational and Institutional Pressures (PR), measured by six items; and
- 5) Professional Training and Experience (TR), measured by five items.

All items were assessed using a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Particular attention was given to auditors' perceived ability to identify significant irregularities in a timely manner, measured within the Auditor Judgment dimension, as the outcome most directly related to the purpose of the present study.

The empirical analysis was conducted in several stages.

First, descriptive statistics were used to summarize the respondents' characteristics and the distribution of responses. The internal consistency of the multi-item constructs was subsequently assessed using Cronbach's alpha. Composite scores were calculated as the mean of the respective items for each construct.

Given the relatively small sample size and the ordinal nature of the Likert-scale data, Spearman's rank-order correlation coefficient (ρ) was used as the primary measure of bivariate association.

The analysis focused particularly on the relationship between professional skepticism and auditors' perceived ability to identify significant irregularities.

3. Results and discussion

Table 1 presents the demographic and professional characteristics of the respondents. The final sample consisted of 35 state auditors, of whom 62.9% were female and 37.1% were male. The respondents represented different age groups and levels of professional experience, with the largest proportion aged between 31 and 40 years (34.3%).

Table 1: Descriptive statistics of the survey respondents

Feature	Description	Female	Male	Total	Percentage
Job category	Category B	5	1	6	17.1%
	Category V	17	12	29	82.9%
	Total	22	13	35	100%
Gender	Female	22	–	22	62.9%
	Male	–	13	13	37.1%
	Total	22	13	35	100%
Age	Up to 30 years	3	3	6	17.1%
	31–40 years	6	6	12	34.3%
	41–50 years	6	2	8	22.9%
	Over 50 years	7	2	9	25.7%
	Total	22	13	35	100%
Audit experience	Up to 5 years	13	7	20	57.1%
	6–10 years	1	4	5	14.3%
	11–15 years	1	1	2	5.7%
	Over 15 years	7	1	8	22.9%
	Total	22	13	35	100%
Professional training in the last 3 years	Yes	22	13	35	100%
	No	0	0	0	0%
	Total	22	13	35	100%
Types of audits*	Financial audit	21	10	31	88.6%
	Compliance audit	14	6	20	57.1%
	Performance audit	14	6	20	57.1%
	IT audit	2	3	5	14.3%
	Other	2	3	5	14.3%

* Multiple responses were allowed for types of audits; therefore, percentages do not sum to 100

Source: Author's research

More than half of the respondents (57.1%) had up to five years of audit experience, while 22.9% reported more than 15 years of experience. In terms of hierarchical position, 82.9% belonged to Category V and 17.1% to Category B. Within the SAO's organizational structure, Category B comprises auditors in managerial and supervisory positions, whereas Category V comprises operational auditors ranging from junior to principal auditor; this distinction was used to explore whether differences in hierarchical position and professional responsibilities are associated with differences in skeptical audit behavior. All respondents had participated in professional training during the previous three years. Financial auditing was the most

frequently reported area of audit experience (88.6%), followed by compliance and performance auditing (57.1% each).

3.1. Descriptive statistics and reliability analysis

Table 2 presents the descriptive statistics and internal consistency coefficients for the five constructs included in the study.

Table 2: Descriptive Statistics and Reliability Analysis

Construct	No. of items	Mean	SD	Cronbach's α
Professional Skepticism (PS)	10	4.629	0.400	0.834
Auditor Judgment (AJ)	4	4.629	0.490	0.903
Organizational Support (OS)	5	4.451	0.743	0.912
Perceived Organizational and Institutional Pressures (PR)	6	2.867	1.030	0.871
Professional Training and Experience (TR)	5	4.589	0.475	0.769

Note: N = 35. Construct scores represent the mean of the respective items measured on a five-point Likert scale.

Source: Author's research

All five constructs demonstrated satisfactory internal consistency, with Cronbach's alpha coefficients ranging from 0.769 to 0.912. Professional Skepticism recorded an alpha coefficient of 0.834, while Auditor Judgment and Organizational Support demonstrated particularly high reliability ($\alpha = 0.903$ and $\alpha = 0.912$, respectively). The reliability coefficients therefore support the use of composite construct scores in the subsequent analyses.

The descriptive results reveal a high reported level of professional skepticism among the surveyed state auditors ($M = 4.629$, $SD = 0.400$). Similarly high scores were observed for auditor judgment ($M = 4.629$), professional training and experience ($M = 4.589$), and organizational support ($M = 4.451$). In contrast, perceived organizational and institutional pressures recorded a substantially lower mean value ($M = 2.867$), indicating that such pressures were generally perceived as less characteristic of respondents' audit environment.

From the perspective of irregularity identification, several individual responses are particularly noteworthy. Respondents reported very strong agreement with seeking additional evidence when explanations are incomplete or inconsistent ($M = 4.829$), remaining alert to possible irregularities and material misstatements ($M = 4.743$), and basing audit conclusions on sufficient and appropriate evidence ($M = 4.800$). Furthermore, the perceived ability to identify significant irregularities in a timely manner recorded a high mean value of 4.571. These findings indicate that skeptical

evidence assessment and attentiveness to irregularities are strongly reflected in the self-reported audit practices of the respondents.

This pattern is particularly relevant from a forensic accounting perspective. Although state auditing and forensic accounting have different institutional mandates, both place considerable importance on questioning evidence, recognizing inconsistencies, and maintaining alertness to circumstances that may indicate irregularities. The descriptive findings therefore provide an initial indication that professional skepticism may represent an important behavioral mechanism supporting irregularity identification in public sector audits.

3.2. Relationships between the main constructs

Spearman's rank-order correlations were calculated to examine the relationships among the five constructs. The results are presented in Table 3.

Table 3: Spearman Correlations Between the Main Constructs

Construct	PS	AJ	OS	PR	TR
PS	1.000				
AJ	0.794***	1.000			
OS	0.612***	0.576**	1.000		
PR	-0.161	-0.201	-0.135	1.000	
TR	0.427*	0.451**	0.436**	-0.202	1.000

Note: N = 35. Spearman's rank-order correlation coefficients are reported. $p < .05$; $p < .01$; $p < .001$.

Source: Author's research

The results reveal a strong positive association between Professional Skepticism and Auditor Judgment ($\rho = 0.794$, $p < 0.001$). This indicates that respondents reporting stronger skeptical behavior also tended to report higher-quality professional judgment. Professional Skepticism was also positively associated with Organizational Support ($\rho = 0.612$, $p < 0.001$) and Professional Training and Experience ($\rho = 0.427$, $p = 0.011$).

Organizational Support was positively associated with Auditor Judgment ($\rho = 0.576$, $p < 0.001$), and Professional Training and Experience ($\rho = 0.436$, $p = 0.009$), while Professional Training and Experience also exhibited a positive relationship with Auditor Judgment ($\rho = 0.451$, $p = 0.007$). These findings suggest that supportive organizational conditions and professional development coexist with stronger skeptical behavior and professional judgment.

In contrast, Perceived Organizational and Institutional Pressures were negatively correlated with both Professional Skepticism ($\rho = -0.161$) and Auditor Judgment ($\rho = -0.201$), but neither relationship was statistically significant. Consequently, the present evidence does not support a statistically reliable association between the overall pressure construct and either professional skepticism or auditor judgment.

This result is worth interpreting alongside the item-level descriptive findings. Respondents acknowledged that short deadlines may negatively affect the thoroughness of audit procedures ($M = 3.886$) and that time pressure may reduce professional skepticism ($M = 3.743$), while reporting considerably lower agreement with statements concerning organizational, political, or institutional influence on audit assessments. Thus, time-related constraints appear to be perceived as more relevant than direct external or institutional pressures.

3.3. Professional skepticism and the identification of significant irregularities

Given the forensic-accounting orientation of the study, particular attention was devoted to the relationship between professional skepticism and respondents' perceived ability to identify significant irregularities in a timely manner. The results are presented in Table 4.

Table 4: Professional Skepticism and Perceived Ability to Identify Significant Irregularities

Variable	Mean	SD	Spearman's ρ with irregularity identification	p-value
Professional Skepticism (PS)	4.629	0.400	0.617	<0.001
Organizational Support (OS)	4.451	0.743	0.410	0.014
Perceived Pressures (PR)	2.867	1.030	-0.028	0.872
Training and Experience (TR)	4.589	0.475	0.322	0.059

Note: Irregularity identification refers to respondents' perceived ability to identify significant irregularities in a timely manner ($M = 4.571$, $SD = 0.608$).

Source: Author's research

A strong positive and statistically significant relationship was found between Professional Skepticism and auditors' perceived ability to identify significant irregularities ($\rho = 0.617$, $p < 0.001$). This represents one of the central findings of the study and indicates that state auditors who report stronger skeptical behavior also tend to report greater confidence in their ability to recognize significant irregularities during audit engagements.

Organizational Support was also positively related to irregularity identification ($\rho = 0.410$, $p = 0.014$). Professional Training and Experience showed a positive but marginally non-significant association ($\rho = 0.322$, $p = 0.059$), whereas Perceived Pressures showed virtually no bivariate association with irregularity identification ($\rho = -0.028$, $p = 0.872$).

3.4. Skeptical behaviors associated with irregularity identification

To explore this relationship in greater detail, the individual Professional Skepticism items were correlated with respondents' perceived ability to identify significant irregularities. Table 5 reports the results.

Table 5: Professional Skepticism Behaviors and Identification of Significant Irregularities

Skeptical behavior	Spearman's ρ	p-value
Avoiding conclusions before considering all relevant information	0.564	<0.001
Actively searching for contradictions or inconsistencies in audit evidence	0.547	<0.001
Considering alternative explanations before reaching a final conclusion	0.531	0.001
Questioning information provided by the audited entity	0.510	0.002
Basing conclusions on sufficient and appropriate audit evidence	0.509	0.002
Verifying assertions made by the audited entity	0.482	0.003
Critically assessing the reliability of audit evidence	0.471	0.004
Seeking additional evidence when explanations are incomplete/inconsistent	0.442	0.008
Remaining alert to possible irregularities and material misstatements	0.408	0.015
Revising the initial audit position when new evidence emerges	0.380	0.024

Notes: N = 35. The dependent item measures respondents' perceived ability to identify significant irregularities in a timely manner.

Source: Author's research

Notably, all ten dimensions of professional skepticism were positively and statistically significantly associated with perceived irregularity identification. The strongest relationships were observed for avoiding premature conclusions ($\rho = 0.564$, $p < 0.001$), actively searching for contradictions or inconsistencies in audit evidence ($\rho = 0.547$, $p < 0.001$), and considering alternative explanations ($\rho = 0.531$, $p = 0.001$).

This finding complements the exploratory hierarchical comparison, in which the only statistically significant difference between Category B and Category V auditors concerned the active search for contradictions and inconsistencies in audit evidence ($p = 0.043$). Higher-level auditors reported stronger agreement with this behavior

($M = 5.000$) than operational-level auditors ($M = 4.138$). Given the unequal group sizes, however, this result should be interpreted cautiously.

3.5. Discussion

Taken together, the findings provide consistent evidence that professional skepticism occupies an important position in state auditors' self-reported approach to identifying irregularities. The respondents reported high levels of professional skepticism and a strong perceived ability to identify significant irregularities, while the correlation analysis demonstrated a strong positive association between these two dimensions.

The results are particularly informative when professional skepticism is viewed through behaviors relevant to forensic accounting. Rather than representing merely a general professional attitude, the findings suggest that its practical relevance lies in specific evidence-evaluation behaviors, particularly withholding premature conclusions, scrutinizing contradictory evidence, and considering alternative explanations. These behaviors may assist auditors in recognizing circumstances requiring additional scrutiny.

Professional training and experience also appear relevant to skeptical audit behavior. Respondents strongly agreed that experience helps them identify inconsistencies and risky situations more easily ($M = 4.771$), that audit experience contributes to more critical evidence assessment ($M = 4.657$), and that professional training assists in recognizing risks and irregularities ($M = 4.486$). At the construct level, Training and Experience was positively associated with both Professional Skepticism ($\rho = 0.427$, $p = 0.011$) and Auditor Judgment ($\rho = 0.451$, $p = 0.007$). Its association with perceived irregularity identification was also positive, although it fell slightly short of conventional statistical significance ($\rho = 0.322$, $p = 0.059$). Taken together, these findings suggest that training and accumulated professional experience form an important part of the broader context in which skeptical judgment is developed and exercised.

Organizational support also demonstrated meaningful relationships with skeptical audit behavior. It was positively associated with Professional Skepticism ($\rho = 0.612$, $p < 0.001$), Auditor Judgment ($\rho = 0.576$, $p < 0.001$), and perceived irregularity identification ($\rho = 0.410$, $p = 0.014$). These findings suggest that an organizational environment that encourages critical thinking, supports additional questioning and evidence gathering, and protects professional judgment may provide favorable conditions for the exercise of professional skepticism and irregularity-sensitive auditing.

Conversely, the overall pressure construct did not demonstrate statistically significant relationships with either Professional Skepticism or irregularity identification. Nevertheless, the item-level findings indicate that respondents distinguish between different forms of pressure. Time-related constraints were perceived as more consequential than organizational or political pressures, suggesting that future research should consider separating time-budget pressure from

broader institutional or external pressure rather than treating them as a single dimension.

The findings extend the relevance of professional skepticism beyond conventional discussions of audit judgment and audit quality. In a public sector environment, skeptical behavior appears closely connected with auditors' perceived capacity to recognize significant irregularities and potentially risky circumstances. This provides an important bridge between public sector auditing and forensic accounting, while respecting the distinction between identifying audit irregularities or warning signs and formally detecting or investigating fraud.

Conclusion

This study examined professional skepticism among state auditors and its relationship with their perceived ability to identify significant irregularities in public sector audits. Based on evidence from auditors employed by the SAO, the study also considered organizational support, professional training and experience, and perceived organizational and institutional pressures as contextual factors surrounding skeptical audit behavior.

The findings demonstrate a high level of professional skepticism among the surveyed state auditors and, more importantly, a strong positive association between professional skepticism and auditors' perceived ability to identify significant irregularities. All individual skeptical behaviors were positively and significantly associated with irregularity identification, with the strongest relationships observed for avoiding premature conclusions, actively searching for contradictions and inconsistencies, and considering alternative explanations. Organizational support was also positively associated with irregularity identification, while professional training and experience showed a positive but marginally non-significant relationship. Perceived organizational and institutional pressures, by contrast, showed no statistically significant association with irregularity identification.

The results contribute to behavioral auditing research by extending the study of professional skepticism to the comparatively underexplored SAI environment and by highlighting its relevance for irregularity-sensitive audit behavior. From a practical perspective, the findings suggest that SAIs should continue to foster an environment that encourages auditors to question evidence, investigate inconsistencies, consider alternative explanations, and seek additional evidence when necessary. Such behaviors may strengthen auditors' ability to recognize circumstances requiring further scrutiny and, ultimately, support the reliability of public sector audit findings. Importantly, identifying irregularities or warning signs should not be equated with formally detecting or investigating fraud.

The findings should nevertheless be interpreted in light of several limitations. The relatively small sample limits the generalizability of the results, while the cross-sectional and self-reported nature of the data does not permit causal conclusions.

Furthermore, the study measures auditors' perceived rather than objectively observed ability to identify significant irregularities. Future research could therefore examine larger samples across multiple SAIs, employ experimental or scenario-based designs, and develop more comprehensive measures of irregularity-identification ability. The findings indicate that professional skepticism represents an important behavioral attribute associated with state auditors' perceived ability to recognize significant irregularities and circumstances requiring additional audit scrutiny.

References

- Agustina, F., Nurkholis, N., & Rusydi, M. K. (2021). Auditors' professional skepticism and fraud detection. *International Journal of Research in Business and Social Science*, 10(4), 275–287. <https://doi.org/10.20525/ijrbs.v10i4.1214>
- Bonollo, E. (2019). Measuring supreme audit institutions' outcomes: Current literature and future insights. *Public Money & Management*, 39(7), 468–477. <https://doi.org/10.1080/09540962.2019.1583887>
- Bozhinovska Lazarevska, Z., Trpeska, M., Atanasovski, A., & Jancevska, I. (2017). The professional scepticism of internal auditors in context of fraud identification – Research in the Republic of Macedonia. Paper presented at the 22nd EBES Conference, Rome, Italy, 24–26 May 2017.
- Budiadnyani, N. P., & Dewi, P. P. R. A. (2024). Professional skepticism and audit judgment: The moderating effect of audit experience. *Journal of Applied Sciences in Accounting, Finance, and Tax*, 7(2), 106–116. <https://doi.org/10.31940/jasafint.v7i2.106-116>
- Butar Butar, S. G. A. (2018). The application of professional skepticism in fraud detection: A case study at BPKP Representative of Central Java Province. *Asia Pacific Fraud Journal*, 2(2), 175–184. <https://doi.org/10.21532/apfj.001.17.02.02.05>
- Cordery, C. J., & Hay, D. (2025). Public sector audit: New Public Management influences and eco-system driven reforms. *Journal of Public Budgeting, Accounting & Financial Management*, 37(4), 595–618. <https://doi.org/10.1108/JPBAFM-07-2023-0124>
- Dionisijev, I., & Bozhinovska Lazarevska, Z. (2024). Enhancing sustainable progress: An analysis of supreme audit institutions' performance audits and information disclosure practices. *Journal of Public Budgeting, Accounting & Financial Management*, 37(2), 254–272. <https://doi.org/10.1108/JPBAFM-06-2023-0092>
- Dionisijev, I., Bozhinovska Lazarevska, Z., Trpeska, M., & Atanasovski, A. (2022). The supreme audit institution in the Republic of North Macedonia – Two decades of building public trust. *Journal of Public Budgeting, Accounting & Financial Management*, 35(4), 515–535. <https://doi.org/10.1108/JPBAFM-08-2021-0129>

- Ghani, E. K., Ilias, A., Muhammad, K., & Mohd Ali, M. (2022). Exploring Hurtt's professional scepticism scale for public sector accountants in Malaysia. *IPN Journal of Research and Practice in Public Sector Accounting and Management*, 12(1), 57–77. <https://doi.org/10.58458/ipnj.v12-01030076>
- Glover, S. M., & Prawitt, D. F. (2014). Enhancing auditor professional skepticism: The professional skepticism continuum. *Current Issues in Auditing*, 8(2), P1–P10. <https://doi.org/10.2308/ciia-50895>
- Grossi, G., Hay, D. C., Kuruppu, C., & Neely, D. (2023). Changing the boundaries of public sector auditing. *Journal of Public Budgeting, Accounting & Financial Management*, 35(4), 417–430. <https://doi.org/10.1108/JPBAFM-05-2023-0079>
- Hay, D., & Cordery, C. (2018). The value of public sector audit: Literature and history. *Journal of Accounting Literature*, 40, 1–15. <https://doi.org/10.1016/j.acclit.2017.11.001>
- Hurtt, R. K. (2010). Development of a scale to measure professional skepticism. *Auditing: A Journal of Practice & Theory*, 29(1), 149–171. <https://doi.org/10.2308/aud.2010.29.1.149>
- Hurtt, R. K., Brown-Liburd, H., Earley, C. E., & Krishnamoorthy, G. (2013). Research on auditor professional skepticism: Literature synthesis and opportunities for future research. *Auditing: A Journal of Practice & Theory*, 32(Supplement 1), 45–97. <https://doi.org/10.2308/ajpt-50361>
- Nelson, M. W. (2009). A model and literature review of professional skepticism in auditing. *Auditing: A Journal of Practice & Theory*, 28(2), 1–34. <https://doi.org/10.2308/aud.2009.28.2.1>
- Quadackers, L., Groot, T., & Wright, A. (2014). Auditors' professional skepticism: Neutrality versus presumptive doubt. *Contemporary Accounting Research*, 31(3), 639–657. <https://doi.org/10.1111/1911-3846.12052>
- Rininda, B. P. (2024). Auditor's professional skepticism and experience on fraud detection: The moderating role of professional ethical commitment. *Jurnal Dinamika Akuntansi*, 16(1). <https://doi.org/10.15294/jda.v16i1.4037>
- Samiolo, R., Spence, C., & Toh, D. (2024). Auditor judgment in the fourth industrial revolution. *Contemporary Accounting Research*, 41(1), 498–528. <https://doi.org/10.1111/1911-3846.12901>
- Sari, R., & Andrian, S. (2023). The influence of professional skepticism, auditor's experiences, and competency on fraud detection. *International Journal of Research in Business and Social Science*, 12(3), 384–390. <https://doi.org/10.20525/ijrbs.v12i3.2569>
- Sayed Hussin, S. A. H., Iskandar, T. M., Saleh, N. M., & Jaffar, R. (2017). Professional skepticism and auditors' assessment of misstatement risks: The

moderating effect of experience and time budget pressure. *Economics & Sociology*, 10(4), 225–250. <https://doi.org/10.14254/2071-789X.2017/10-4/17>

Shirowzhan, H., & Fakhari, H. (2024). Clarifying the decision-making mystery: Drivers of professional skepticism, ego depletion and overconfidence in independent auditors' quality of judgment. *Managerial Auditing Journal*, 39(7), 821–842. <https://doi.org/10.1108/MAJ-04-2024-4317>

Sumartono, S., Ermawati, Y., & Rorong, T. M. (2022). Influence of experience and professional skepticism on audit judgement of government external auditors. *Jurnal Akuntansi dan Auditing Indonesia*, 26(2), 211–223. <https://doi.org/10.20885/jaai.vol26.iss2.art10>

Volodina, T., & Grossi, G. (2025). Digital transformation in public sector auditing: Between hope and fear. *Public Management Review*, 27(5), 1444–1468. <https://doi.org/10.1080/14719037.2024.2402346>

Appendix

Survey Questionnaire – Likert-Scale Items

A. Professional Skepticism (PS)

PS1. I critically assess the reliability of audit evidence obtained during an audit.

PS2. I frequently question information provided by the audited entity.

PS3. I seek additional audit evidence when explanations provided by the audited entity are incomplete or inconsistent.

PS4. I rarely accept assertions made by the audited entity without additional verification.

PS5. I actively look for contradictions or inconsistencies in audit evidence.

PS6. I avoid reaching conclusions before considering all relevant information.

PS7. I remain alert to possible irregularities and material misstatements throughout the audit process.

PS8. I am willing to revise my initial position when new audit evidence becomes available.

PS9. I base my audit conclusions on sufficient and appropriate audit evidence.

PS10. I consider alternative explanations before reaching a final audit conclusion.

B. Auditor Judgment (AJ)

AJ1. My audit assessments and conclusions are consistent with applicable auditing standards and methodology.

AJ2. I believe I can identify significant irregularities in a timely manner during an audit.

AJ3. I am able to objectively assess audit findings even when working under pressure.

AJ4. I am able to appropriately assess audit risk in complex audit situations.

C. Organizational Support (OS)

OS1. My institution encourages critical thinking and professional skepticism during audit engagements.

OS2. My managers support auditors when they ask additional questions or request additional audit evidence.

OS3. Professional training provided by my institution contributes to the development of professional skepticism.

OS4. The audit methodology and procedures applied by my institution support independent professional judgment.

OS5. The organizational culture within my institution encourages the application of professional skepticism.

D. Perceived Organizational and Institutional Pressures (PR)

PR1. Short audit deadlines may negatively affect the thoroughness of audit procedures.

PR2. Time pressure may reduce the level of professional skepticism applied during an audit.

PR3. Time constraints sometimes lead auditors to complete audit procedures without obtaining sufficient additional evidence.

PR4. Organizational or external pressures may influence auditors' professional assessments.

PR5. Sensitive audit situations may discourage auditors from asking additional questions or pursuing further evidence.

PR6. Political or institutional sensitivity may influence professional judgment during an audit.

E. Professional Training and Experience (TR)

TR1. Professional training helps me identify audit risks and potential irregularities more effectively.

TR2. My audit experience contributes to a more critical assessment of audit evidence.

TR3. Over time, I have become more skeptical when evaluating information provided during audit engagements.

TR4. Continuous professional development is important for maintaining high-quality professional judgment.

TR5. My audit experience helps me identify inconsistencies and risky situations more easily.

Note: All statements were measured on a five-point Likert scale ranging from 1 (Strongly disagree) to 5 (Strongly agree).

Savka Vučković Milutinović
University of Belgrade, Faculty of
Economics and Business
ORCID: 0000-0003-1065-0683

Tijana Ostojić
University of Belgrade, Faculty of
Economics and Business
ORCID: 0009-0006-3562-6152

INTERNAL AUDIT AS A VALUABLE ACTOR IN ADDRESSING FRAUD: WHERE DOES IT STAND IN SERBIA?

UDK 657.6:343.72

DOI <https://doi.org/10.54318/czf.2026.svm.005>

Abstract

Fraud risk represents one of the traditional, yet still highly relevant, areas of risk management in organisations. The role of internal audit in reducing organisations' exposure to fraud has long been recognised, and its importance has further increased alongside growing expectations regarding corporate governance, risk management, and internal control. The professional framework recognises internal audit as an important actor in assessing and responding to fraud risk, although the boundaries of its responsibilities in this area are not always clearly defined. At the same time, internal auditors are expected to possess appropriate competencies and to continuously consider fraud risk when planning and performing audit engagements.

This paper is focused on the role of internal audit in addressing fraud in organisations in Serbia. Specifically, it explores fraud-related activities in which internal audit participates, as well as the extent to which organisational and professional conditions may facilitate or constrain its contribution. Insights into current practice are provided through the lens of internal auditors, given the limited availability of organisation-level data on the characteristics of internal audit functions. The findings indicate that internal audit contributes to a range of fraud-related activities, although the intensity of its engagement varies, with fraud prevention emerging as its dominant role. A higher level of engagement is associated with more developed internal auditors' competencies, while certain organisational and professional conditions are also proved to be influential. The results could be useful for managers and internal auditors by suggesting avenues to work on in strengthening anti-fraud efforts of internal auditing.

Key words: internal audit, fraud, internal audit competencies, organisational conditions, greenwashing

JEL classification: M42, M14, K13

Introduction

The risk environment in which organisations operate is constantly evolving, requiring them to continuously reassess and adapt their risk management practices. Fraud risk is the traditional one, but the methods and opportunities for fraudulent activities continue to develop alongside technological, organisational, and regulatory changes. According to PwC's (2024) global survey among companies, fraud and corruption remain pervasive and challenging issues. The findings show that procurement fraud, in particular, is one of the three most disruptive forms of economic crime experienced by companies globally over the previous two years. Moreover, vast majority of respondents believe that corruption risks are either increasing or remaining unchanged. One estimate suggests that organisations lose, on average, 5% of their revenues to occupational fraud (ACFE, 2026). In addition, compliance risks related to anti-corruption, anti-money laundering, and fraud have also gained prominence, ranking among the top five compliance risk priorities (PwC, 2025).

Since the primary function of internal audit (IA) is to strengthen an organisation's corporate governance, risk management, and control (IIA, 2024, p. 7), examining the competencies and roles of IA that may contribute to combating fraud and corruption is warranted. Empirical research indicates that IA activity related to fraud risk varies significantly across companies (Bonrath & Eulerich, 2024). In addition, responsibilities for addressing fraud risk remain somewhat ambiguous, reflecting broader debates about internal governance structures and the transition from the Three Lines of Defense framework to the Three Lines Model (Eulerich M. , 2021). Nevertheless, expectations regarding internal audit's role in fraud prevention and detection remain high, with evidence suggests that IA is the second most effective mechanism for detecting occupational fraud, after tips (ACFE, 2026).

Designing an effective IA is a complex and challenging task that requires a thorough understanding of the factors that drive its contribution to organisational performance (Ma'ayan & Carmeli, 2016). This is particularly relevant when developing an IA function capable of effectively addressing fraud risk. Nevertheless, identifying and leveraging the potential of IA to contribute to anti-fraud efforts can provide multiple benefits to organisations and should therefore be actively pursued. In our study, we respond to the call for further research on the role of IA in addressing fraud risk across different country settings (Bonrath & Eulerich, 2024). Specifically, we examine IA practices in this area in Serbia, where the IA profession has a relatively shorter tradition than in many developed countries.

Our findings contribute to a better understanding of the commonalities and differences in IA practices across developing and developed country contexts. The results also shed light on the determinants that influence the IA engagement related to fraud in a positive or negative way, while identifying areas where IA practices could be further improved. As sustainability has become an increasingly important

issue for businesses worldwide, we also consider a specific type of fraud risk-greenwashing risk and examine the role that IA may play in addressing it.

The remainder of the paper is structured as follows. The first section discusses internal auditors' professional perspective in relation to fraud, drawing on the IA professional body of knowledge, including mandatory requirements. The second section presents literature review and research questions. Section three outlines the research sample and methodology, while Section four presents and discusses the results. Finally, the concluding section highlights the key findings and implications of the study.

1. The professional perspective on IA's role in addressing fraud

IA has evolved into an integral component of corporate governance in many organisations, particularly public-interest entities, alongside the strengthening of its professional identity. However, the IA function must continuously demonstrate its value to the organisation. In their analysis of voluntary IA adoption among NASDAQ-listed companies, Watts et al. (2026) point out that not all companies consider investment in IA appropriate. Given their limited resources, companies must weigh IA against other competing priorities (p. 427). Among the various contributions of IA, the protection of organisations against fraud represents an important area of value. This role has become particularly prominent as fraud has gained greater visibility. In this context, the wave of financial scandals in US at the beginning of XXI century pushed the demand for IA given its role in assessing and strengthening internal control. Carcello et al. (2005) confirm empirically on the sample of US public companies that IA budgets and staffing levels, as well as meetings with the audit committee and meeting length, increased considerable from 2001 to 2002.

The IA profession has responded to growing stakeholder expectations regarding its responsibilities in relation to fraud. However, these expectations may at times exceed the role that IA performs. A survey of institutional investors conducted by CAQ & KRC Research (2025) indicates that respondents perceive IA as having the primary responsibility for preventing and detecting fraud within companies, ranking it above company management and boards of directors. Such a perception is not consistent with the professional framework, according to which primary responsibility for preventing and detecting fraud rests with management, under the oversight of the board of directors, since management is responsible for establishing and maintaining an effective system of internal controls and for managing fraud risk (COSO & ACFE, 2023; IIA, 2024). In contrast, internal audit should support the organisations' anti-fraud efforts by providing assurance over controls designed to prevent and detect fraud (IIA, 2019). Therefore, clarifying the position of internal audit in relation to fraud remains an important task for the profession, so that internal auditors receive clearer guidance and stakeholders develop more realistic expectations. In

their study, DeZoort and Harrison (2008) show that professional obligation has the highest influence on internal auditors' perceive responsibility for fraud detection.

The IIA Global Internal Audit Standards address fraud through several principles and requirements. Table 1 presents the relevant parts of the Standards.

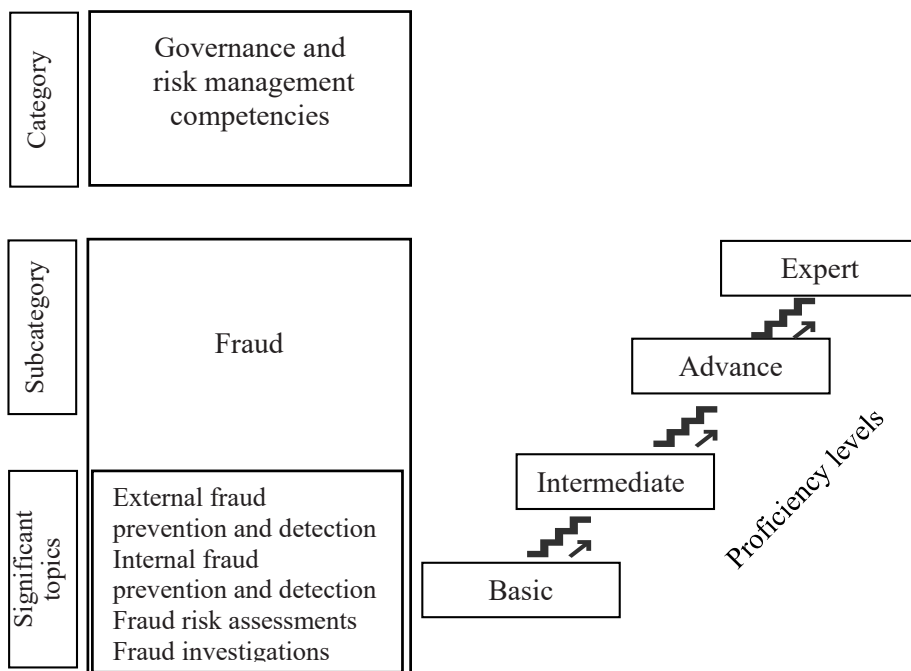
Table 1: Relevant IIA Standards addressing fraud

Principle (P)	Standard (S)
P 3 Demonstrate Competency	S 3.1 Competency
P 4 Exercise Due Professional Care	S 4.2 Due Professional Care
P 9 Plan Strategically	S 9.4 Internal Audit Plan
P 11 Communicate Effectively	S 11.5 Communicating the Acceptance of Risks
P 13 Plan Engagements Effectively	S 13.2 Engagement Risk Assessment

Source: Authors' presentation based on (IIA, 2024)

For internal auditor to satisfactory perform their roles, competencies are crucial. Therefore, it is warrant for the profession to conduct analysis of needed competencies. The current requirement for competencies in Standard 3.1 encompasses different types of knowledge, skills and abilities. Among other, fraud-focused competencies are explicitly stated. They could be further examined by using IIA Competency Framework. IA profession's approach to the competency domains is not static; rather evolving in response to emerging challenges and reflected in revised competency frameworks. However, fraud-focused competencies are included in continuity. Figure 1 presents the position and structure of these competencies within contemporary IIA framework.

It is suggested by framework (IIA, 2025) that basic level competencies are required for entry-level and staff-level internal auditors and cover understanding of fraud concepts, recognising the role of IA related to fraud and assisting in activities considering fraud risk. However, progress in proficiency means ability for conducting tasks relevant for intermediate level, such as conducting audit procedures to identify and analyze fraud indicators as well as assessing fraud prevention and detection controls and determine discrepancy. The advanced level includes further progress so that one is able to lead fraud risk-related engagement and assess governance and risk management processes in overseeing and mitigating fraud risks. Finally, expert level is the highest one and includes developing advanced fraud detection methodology, providing leadership on fraud risk management, advising board and management on approaches to governance and risk management related to fraud.

Figure 1: Fraud-focused competencies within IIA competency framework

Source: Authors' presentation based on IIA competency framework (IIA, 2025)

In addition to the aforementioned competencies, there are also fraud-related competencies identified within several other competency categories (subcategories), such as IA competencies (audit methodology) and professional competencies (data analysis). Certain operational areas are particularly vulnerable to fraud, notably accounting, sales, and supply chain management. This vulnerability is reflected in the inclusion of specific fraud-related competencies within the respective competency subcategories.

Overall, the profession's perspective on competencies, as articulated through the competency framework, encourages internal auditors to strengthen their role in organisations' efforts against frauds. Furthermore, as Table 1 shows, it is the professional duty of internal auditors to exercise due professional care when providing services, what includes assessing the probability of fraud. This contributes to internal auditors' ongoing vigilance for fraud indicators. Beyond this, fraud risk must be considered when developing the audit plan as part of strategic planning, as well as during engagement planning, with the assessment of specific fraud risks forming a part of the risk assessment process. Finally, the standard on communication stipulates the reporting of unacceptable levels of risk, while determining whether a risk exceeds the established risk appetite requires consideration of its potential impact, including the occurrence of fraud. Taken

together, these requirements set out in the standards demonstrate that consideration of fraud is integral to virtually all aspects of internal auditors' work.

Such an approach in the IA standards provides the foundation upon which the role of IA in addressing fraud has been shaped. The IIA, as well as other professional associations, recognises IA as playing a "critical role in helping organisations assess, manage, and respond to fraud risk" (ACF, 2026). Hence, organisations could prevent, detect and respond to fraud more effectively with IA. However, some limitations are also emphasised. It is not generally expected that fraud investigation is part of the IA's role, although it could be the case if it is included in the IA mandate in organisation and IA has the proper skill sets (IIA, 2019).

A review of external auditing standards, specifically International Standard on Auditing 240 (ISA 240) - The Auditor's Responsibilities Relating to Fraud in an Audit of Financial Statements, reveals that an effective IA function is regarded as a factor that constrains improper conduct by management while, conversely, high staff turnover or ineffective personnel within the IA function is viewed as a fraud risk factor associated with the opportunity condition for fraudulent financial reporting. Accordingly, external auditors are interested in understanding the objectives of the IA function with regard to monitoring anti-fraud controls, the procedures performed in this area, and the IA function's knowledge of fraud and control deficiencies, as well as management's response to any findings identified. This implies that the final contribution of IA in the anti-fraud efforts cannot be taken for granted and must instead be assessed in light of a range of relevant factors.

2. Literature review and research questions

Fraud is globally ranked by IA leaders among the six highest audit priorities, where their IA function spend the most time and effort (Internal Audit Foundation, 2025). Although such focus of IA could be viewed as intrinsically motivated by the need of both the organisations and IA functions to avoid the adverse consequences of fraudulent activities, external regulatory-type pressures also play the role. In the Italian context, the adoption of corporate fraud law constituted a pressure factor for companies to introduce or broaden the risk assessment activities by implementing a specific fraud risk assessment project (Allegrini & D'onza, 2003, p. 197). Also, the wave of fraudulent financial statements in US influenced the New York Stock Exchange (NYSE) to introduce in 2003 a requirement for listed companies to maintain an internal audit function.

Many research studies confirm empirically the capacity of IA in fraud risk management. In the context of fraud triangle theory, the impact of IA is commonly associated with the opportunity condition, since IA assesses and improves internal control, including fraud prevention and detection controls. However, IA can also reduce organisations' exposure to the adverse impacts of the other two conditions, namely pressures and rationalizations, by fostering an ethical culture within

organisations, that influences individuals' behaviours towards ethical-decision making and away from committing fraud.

Asare et al. (2008) argue that internal auditors warrant confidence in their capacity to play a critical role in their organisations' anti-fraud and risk management programs. Their research suggests that internal auditors are capable of adopting an "external lens" when assessing fraud risk factors, including those that may be particularly sensitive given their organisational position, such as management incentives and the quality of audit committee.

IA demonstrates its fraud-related roles in different sector settings. Nadirsyah et al. (2024) provide empirical evidence that the IA contributes positively to fraud prevention in public-sector organisations. Their findings further indicate that fraud prevention represents one of the mechanisms through which IA contributes to effective governance. In addition, IA contributes to stemming different types of fraud. Asiedu and Deffor (2017) find that an effective IA function is associated with lower levels of administrative corruption in the Ghanaian public sector. However, institutional conditions and organisational characteristics play an important role in enabling IA to operate effectively. Novatiani et. al (2024) examine fraud in financial statements of state-owned enterprises and show evidence that quality IA positively effects prevention of this type of fraud, as well as mediates the effects of internal control on fraudulent financial statements prevention.

De Zoort and Harrison (2018) highlight a distinction between auditors' prescribed and perceived responsibilities for fraud detection and a need for examination of perceived responsibilities. They find that professional obligation and personal control are significantly associated with auditors' sense of responsibility for fraud detection, and that internal auditors perceive similar levels of responsibilities across all types of occupational fraud, whereas external auditors perceive the greatest responsibility for detecting financial statement fraud. One earlier study also shows that internal auditors perceive their role in preventing, detecting and reporting fraud positively, and moreover they tend to be proactive in searching incidents of fraud (Halbouni, 2015). In addition, Hazami-Ammar (2019) provides empirical support for viewing IA as a function that can play an active role in fraud investigations, rather than being limited to fraud prevention and detection.

Considering the capacity of IA to contribute to organisations' efforts to deal with fraud risk by assuming different roles, as well as the possibility that this capacity may remain underexploited, this study aims to shed light on the anti-fraud engagement of IA in companies operating in Serbia. IA does not have a long tradition in Serbia, as it does in some developed countries, which may mean that more challenging roles, such as those related to fraud, have not yet been fully developed. On the other hand, developing-country settings are often characterised by weaker institutions and rule of law, as well as less-developed organisational practices, potentially increasing organisations' vulnerability to fraud. The involvement of IA

in addressing fraud is therefore of great importance. Against this background, we pose the first research question:

RQ1: What roles does internal audit assume in addressing fraud in organisations operating in Serbia?

Since only competent internal auditors could fulfill their roles, there is need for developing their competencies for anti-fraud activities, what can be done by following framework created by IIA. Empirical evidence also illustrates the focus on these competencies. A large-scale survey of IA professionals conducted by Vitalis et al. (2024) in 2021 reveals that, on average, competencies related to fraud are perceived as very important for internal auditors in the current environment, but as even more important ten years into the future. Furthermore, more than two-thirds of respondents (68.6%) rated fraud-focused competencies as very important or even essential and expected their importance to either remain at the same level or increase in the future. Jokipii & Rautiainen (2024) were also interested in the future of IA and identify three key driving forces for its development: prolific data and its application, globalization and new value chains between organisations. It could be argued that fraud-focused competencies are influenced by all three forces, particularly by the need for improved data analysis to leverage prolific data and identify unusual patterns that may indicate fraudulent activity. At the same time, globalization and increasingly complex relationships across value chains may create greater opportunities for fraud and, consequently, increase the need for fraud-focused competencies within IA. Drouglas et. al (2017) find that training of internal auditors affects positively and significantly the detection of fraud.

Accordingly, this study considers both the level of internal auditors' anti-fraud competencies and the extent to which these competencies are applied in practice as potentially important determinants of IA engagement in addressing fraud risk. However, competencies constitute only one part of a broader set of professional and organisational conditions that may facilitate or constrain IAs' contribution.

The literature suggests that a number of factors may influence the propensity of organisations to establish IA (Watts, Elder, & Hyman, 2026). In addition, established IA could be more or less engaged related to fraud risk (Bonrath & Eulerich, 2024). It is therefore worthwhile to examine the factors that drive IA functions to assume a more advanced role in addressing fraud risk. Based on a sample of IA functions in Germany, Austria, and Switzerland, Bonrath & Eulerich (2024) find that a strong governance environment and the use of technology-based auditing techniques positively influence the IA's role in fraud prevention and detection. In addition, the authors show that this role is strengthened by increased IA involvement with management, while greater interaction with the audit committee is associated with the opposite effect. Khelil et. al (2018) point that moral courage of internal auditors is crucial for their dealing with fraud, and provide evidence about factors fostering moral courage, including positive states of internal auditors (e. g. self-efficacy and resilience), perceived supervisor support and the independence of IA function.

Another study confirms that IA effectiveness, internal auditors' responsibilities and training affect positively and significantly the detection of fraud (Drogalas, Pazarskis, Anagnostopoulou, & Papachristou, 2017). On the other hand, some factors present obstacles for IA to take anti-fraud roles. Smaller and resource constrained firms are even less likely to establish IA (Watts, Elder, & Hyman, 2026) and managers with dark triad traits seem to be able to manipulate internal personnel of IA, so making IA ineffective in curtailing their negative behaviour (Mutschmann, Hasso, & Pelster, 2022). Hence, we define our second research question in the following manner:

RQ2: What factors facilitate or constrain IA engagement in addressing fraud in organisations in Serbia?

The growing interest in sustainability matters leads companies to increasingly report about their environmental, social and governmental (ESG) performance (KPMG, 2024). However, risk of fraudulent statements in sustainability reports has become an important concern and may manifest itself in many different forms (ACFE & Grant Thornton, 2022). Siano et. al (2017) emphasize greenwashing and warn that it could lead not only to decoupling or attention deflection, but also to substantive actions that are in contrast to the sustainability claims. However, Ruiz-Blanco et. al (2022) show that greenwashing was less prevalent among companies in environmentally sensitive industries and in those with external assurance of their sustainability reports. IIA clearly recognises the IA's role in ESG reporting, through assurance and advisory services (IIA, 2021). Still, Rakipi & D'Onza (2023) provide empirical evidence that maturity of companies' ESG practices and stakeholders' salience drive the type of IA's involvement in ESG, ranging from limited to more developed one. Interestingly, Eulerich et al. (2022) find a significant positive association between ESG report assurance and companies' assessment of the environmental pillar as highly relevant, but assurance in ESG field by IA was scarce in comparison with external audit. In their research of social and environmental report assurance, Jones & Solomon (2010) show that half of the respondents expressed a positive attitude toward internal assurance, perceiving it as sufficient. Taken together, these arguments motivate the following research question:

RQ3: How does internal audit contribute to addressing greenwashing risk in organisations in Serbia?

3. Research design

To examine the research questions, we conducted a survey of internal auditors in Serbia. The choice of survey-based approach was motivated by its widespread use in previous empirical research examining the relationship between IA and fraud, including studies based on samples of internal auditors (Burnaby et al., 2011) and chief audit executives (Bonrath & Eulerich, 2024). Drogalas et al. (2017) employed a questionnaire targeting a broader samples of respondents from listed companies, including internal auditors and managers. Since internal audit data are rarely publicly

available, a survey of internal auditors provides a valuable means of gaining insight into the actual practices and experiences of internal auditors.

The questionnaire was developed using Google Forms, and responses were collected during August and September 2026. Invitations to participate in the study were distributed through multiple channels, including publicly available email addresses of professionals available in the Register of Authorized Internal Auditors, maintained by the Chamber of Authorized Auditors of Serbia, and members of the Institute of Internal Auditors of Serbia (IIA Serbia). The use of multiple distribution channels prevented us from reliably determining the response rate. Finally, a total of 55 internal auditors completed the questionnaire.

The first part of the questionnaire was focused on the characteristics of the respondents and the organisations in which they were employed, as summarised in Table 2. The study included internal auditors employed in organisations of varying sizes and operating across different sectors. Most respondents held either the national Authorised Internal Auditor qualification or the internationally recognised Certified Internal Auditor (CIA) qualification. Some respondents also reported other related professional qualifications, such as Certified Fraud Examiner (CFE) and Certified Accountant, as well as certificates in public procurement and compliance.

Table 2: Survey respondents and their firms

Industry sector	
Real sector	30 (54.55%)
Financial sector	25 (45.45%)
Sector by ownership	
Private sector	32 (58.18%)
Public sector	23 (41.82%)
Professional qualifications in internal auditing	
Certified Internal Auditor- Chamber of Authorized Auditors of Serbia	24 (43.64%)
CIA – Certified Internal Auditor (IIA)	14 (25.45%)
Certified Internal Auditor in the Public Sector (Ministry of Finance)	9 (16.36%)
No professional qualification in internal auditing	13 (23.64%)
Other professional qualifications	
Certified Fraud Examiner	5 (9.09%)
Certified Accountant	4 (7.27%)
Other	12 (21.82%)

Note: Since the question allowed multiple responses, the total exceeds 100 percent.

Source: Authors' presentation

The questionnaire was further structured in accordance with the research questions and included various types of questions, such as single-choice, multiple-choice, and open-ended questions. Most questions were mandatory. Respondents were first asked to indicate the roles that IA in their organisations assume in addressing fraud, based on main components of organisations' anti-fraud efforts, namely fraud prevention, detection, investigation and remediation. They were then asked to assess the extent to which IA in their organisations was engaged in specific fraud-related activities, which could be in the assurance or consulting field. A five-point Likert scale was used to assess the extent of IA involvement in individual activities, with 1 indicating no involvement at all and 5 indicating involvement to a very great extent. The next part of the questionnaire comprises questions related to various factors influenced IA anti-fraud contribution. Respondents show the level of agreement (1 = strongly disagree; 5 = strongly agree) with statements dealing with certain organisational and professional factors. In addition, potential barriers were assessed according to the extent to which they constrained IA engagement in this field (1 = not a barrier at all; 5 = a very significant barrier). For respondents from organisations that report on sustainability, additional questions addressed greenwashing risk and the role of IA in this area.

4. Results and discussion

The role of IA may extend across all components of anti-fraud efforts, but the results indicate that the intensity of its engagement is not evenly distributed. The largest proportion of respondents reported IA involvement in fraud detection (83.64%) and prevention (78.18%), while its involvement was considerably less common in fraud investigation and particularly in remediation (Table 3). However, although detection was the most frequently reported area of involvement, when respondents were asked to identify a single dominant component, more than 60% selected prevention, compared with only approximately 22% who selected detection.

Table 3: Internal audit involvement across anti-fraud components

Component	N	%
Prevention	43	78.18
Detection	46	83.64
Investigation	27	49.09
Remediation	12	21.82

Note: Since the question allowed multiple responses, the total exceeds 100 percent.

Source: Authors' presentation

Regarding the type of internal audit services provided, the majority of respondents reported that IA was primarily involved in addressing fraud through assurance engagements (65.45%), while slightly more than one-fifth highlighted its advisory role. This finding is particularly interesting given that Vitalis et al. (2024) highlight

a shift in IA from its traditional compliance-oriented role towards a more consultative one. However, our results suggest that this shift is not yet evident, at least when it comes to addressing fraud risk.

The subsequent survey questions sought to examine the extent of IA involvement in specific fraud-related activities (Table 4). The development of some survey questions was informed by Anderson et al.'s (2017) discussion of the opportunities for internal auditors to provide management with insights into fraud risk. The highest levels of involvement were reported for assessing anti-fraud controls (mean = 3.98) and providing recommendations for their improvement (mean = 3.96), while involvement in employee training, which also represents an important fraud prevention mechanism, was considerably lower (mean = 2.56). A similar disparity was observed in fraud detection, where IA was relatively less involved in developing processes and tools for early fraud detection (mean = 2.65) and whistleblowing procedures (mean = 2.49). The relatively low level of involvement in conducting internal investigations (mean = 2.62) further indicates that IA plays a less prominent role in activities undertaken after fraud has been detected, which may reflect the division of responsibilities between IA and other relevant functions within the organisation.

Table 4: Level of internal audit involvement in fraud-related activities

Activity	Mean	Std. Deviation
Assessment of anti-fraud controls	3.98	1.240
Providing recommendations for improving anti-fraud controls	3.96	1.122
Comprehensive fraud risk assessment	3.67	1.277
Monitoring the implementation of corrective measures following identified fraud	3.00	1.427
Supporting management in developing an ethical culture and a low tolerance for fraud	2.98	1.312
Analyzing the causes of detected fraud and providing recommendations to prevent recurrence	2.91	1.405
Developing processes and tools for early fraud detection	2.65	1.174
Conducting internal fraud investigations	2.62	1.367
Employee training and raising awareness of fraud risk	2.56	1.358
Supporting the development of whistleblowing procedures	2.49	1.345

Source: Authors' presentation

The relatively high standard deviations indicate considerable variation in respondents' assessments of the extent of IA involvement in individual activities. To

examine whether the extent of IA involvement in specific activities differed across sectors, the Mann–Whitney U test was employed, given that involvement in each of the ten activities was measured on a five-point Likert scale and that two independent groups of respondents were compared in each case. However, since ten separate tests were conducted for each grouping, Holm’s correction for multiple comparisons was additionally applied.

When comparing organisations in the non-financial and financial sectors, a statistically significant difference at the 5% significance level was found for providing recommendations to improve anti-fraud controls ($U = 266.5$; $Z = -1.973$; $p = 0.048$). However, this difference was no longer statistically significant after applying Holm’s correction (adjusted $p = 0.480$). Different results emerged when comparing organisations in the private and public sectors. A statistically significant difference was found in respondents’ assessments of IA involvement in supporting the development of whistleblowing procedures ($U = 229.5$; $Z = -2.443$; $p = 0.015$), with internal auditors from the private sector having a higher mean rank than those from the public sector (32.33 vs. 21.98). An even more pronounced difference was observed for conducting internal fraud investigations ($U = 148.0$; $Z = -3.861$; $p < 0.001$), with a mean rank of 34.88 in the private sector compared with 18.43 in the public sector. However, after applying Holm’s correction for multiple comparisons, statistical significance was retained only for conducting internal fraud investigations.

Building on the previous findings regarding the general characteristics of organisations, the question arises as to whether the specific conditions under which IA operates are more relevant to its contribution to anti-fraud efforts. This question is also supported by previous research suggesting that organisational status, competencies, resources, and management support influence IA’s contribution in this area (Kabuye et al., 2017; Bonrath & Eulerich, 2024). According to the results presented in Table 5, regarding the conditions under which IA operates in Serbia, the highest mean score was reported for unrestricted access to data, documentation, systems, and employees necessary for investigating suspected fraud (mean = 3.78), while the availability of technological resources and tools for identifying fraud risks received the lowest mean score (mean = 2.82).

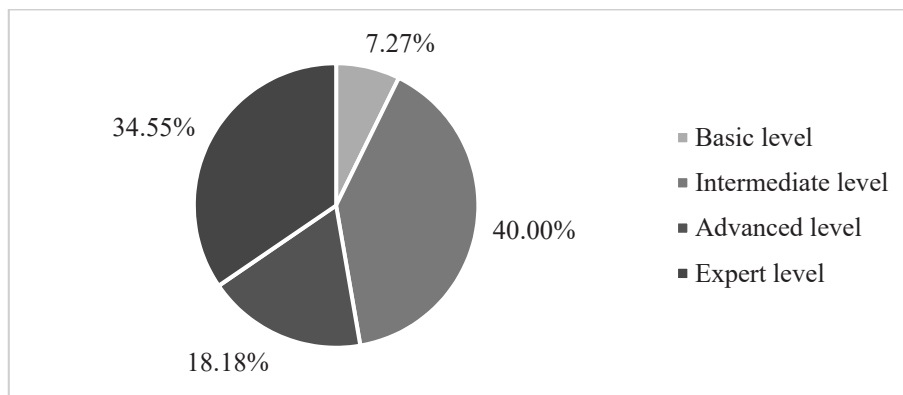
Table 5: Organisational and professional factors influencing internal audit involvement in combating fraud

Factors	Mean	Std. Deviation
Fraud-related mandate	3.45	1.303
Fraud-related competencies	3.45	1.274
Technological resources	2.82	1.140
Access to information	3.78	1.182
Management responsiveness	3.55	1.274

Source: Authors’ presentation

To examine whether these factors are associated with the actual extent of IA's role in fight against fraud, Spearman's rank correlation was employed, with particular attention given to two factors directly related to IA's ability to fulfil its role: the existence of a mandate to address fraud risk and management's response to IA recommendations. The results indicate that the perception of a clearer mandate to address fraud risk is positively associated with IA involvement in assessing anti-fraud controls ($\rho = 0.397$; $p = 0.003$) and providing recommendations for their improvement ($\rho = 0.476$; $p < 0.001$). It is worth recalling that both activities were previously identified as the most prevalent forms of IA engagement in the sample. In contrast, the absence of a statistically significant association between the mandate and the conduct of internal investigations suggests that formal authority to address fraud risk is not, in itself, sufficient for IA to assume an investigative role. When it comes to the second factor, the results reveal a relatively strong positive association between an adequate management response and IA involvement in providing recommendations to improve anti-fraud controls ($\rho = 0.588$; $p < 0.001$), as well as supporting management in developing an ethical culture and low tolerance for fraud ($\rho = 0.501$; $p < 0.001$). While Bonrath & Eulerich (2024) found a positive association between the intensity of IA interactions with management and its engagement in fraud prevention and detection, our study examines this relationship from a different perspective, showing that a more favourable management response is associated with greater IA involvement in providing recommendations and fostering an ethical culture.

As previously noted, fraud-related competencies were assessed in accordance with the IIA Internal Audit Competency Framework (2025). The largest proportion of respondents rated their competencies at the advanced level (40.00%), while slightly more than half reported competencies at the basic or intermediate level (Figure 2). Regarding fraud-related training, 38.18% of respondents reported attending such training occasionally, while only 5.45% did so very frequently. At the same time, 27.27% attended training very rarely, and 7.27% never attended it. A statistically significant positive correlation was found between the frequency of fraud-related training and the level of competencies ($\rho = 0.451$; $p < 0.001$). Furthermore, a positive and statistically significant correlation was found between the level of fraud-related competencies and the overall extent of IA engagement in the activities examined ($\rho = 0.361$; $p = 0.007$). It should be noted that the composite measure of overall IA engagement in fraud-related activities was calculated by averaging the ratings across the ten activities included in the questionnaire. The use of this composite measure was supported by the high internal consistency of the ten items (Cronbach's $\alpha = 0.902$).

Figure 2: Self-assessed competency levels of internal auditors

Source: Authors' presentation

In addition to assessing their level of competencies, respondents were asked to indicate the extent to which they consider fraud risk when performing audit engagements. The results show that internal auditors pay considerable attention to this issue, with a mean score of 4.44 and 89.09% of respondents assigning a rating of 4 or 5. To examine whether greater attention to fraud risk is associated with the conditions under which IA operates, a composite measure of organisational and professional factors was constructed by averaging respondents' ratings across five statements (Cronbach's $\alpha = 0.819$). Spearman's rank correlation revealed a positive and statistically significant association between the attention paid to fraud risk and this composite measure ($\rho = 0.281$; $p = 0.038$). Similarly, Kabuye et al. (2017) show that internal auditors are more likely to perform adequate fraud management activities when the IA function has an appropriate organisational status and competencies. Looking at each factor separately, positive and statistically significant correlations were found between the attention paid to fraud risk and the availability of appropriate technological resources and tools ($\rho = 0.361$; $p = 0.007$), as well as between the attention paid to fraud risk and internal auditors' knowledge and competencies in this area ($\rho = 0.341$; $p = 0.011$). The correlation with technological resources is further supported by the findings of Bonrath & Eulerich (2024), who found that the use of technology-based approaches to risk identification is associated with a greater IA focus on fraud prevention and detection.

The study also examined potential barriers to IA's contribution to anti-fraud efforts, with respondents' perceptions presented in Table 6. Among the barriers examined, the highest mean scores were recorded for prioritizing other IA activities (mean = 3.15) and insufficiently developed technology (mean = 3.11). While the second finding is consistent with the previously reported relatively low rating of the availability of appropriate technological resources and tools, the first suggests that lower levels of IA engagement may reflect a gap between the importance attached to fraud risk and the resources available to IA to address it alongside other priorities.

Table 6: Perceived barriers to internal audit involvement in fraud-related activities

Barrier	Mean	Std. Deviation
Insufficient number of internal audit staff	2.75	1.493
Priority given to other internal audit activities	3.15	1.367
Lack of specific fraud-related knowledge and skills	2.87	1.248
Insufficiently developed technology	3.11	1.257
Lack of management support	2.73	1.340
Unclear internal audit responsibilities regarding fraud	2.87	1.402
Risk of pressure or negative consequences	2.55	1.385
Insufficient cooperation with other relevant functions	2.47	1.345

Source: Authors' presentation

Although the risk of pressure or negative consequences was not among the most prominent barriers in the overall sample (mean = 2.55), the analysis by organisational ownership revealed significant differences in its perception between the public and private sectors. Internal auditors in the public sector perceived this barrier as more significant than their counterparts in the private sector, with the difference remaining statistically significant after correction for multiple comparisons ($U = 153.00$; $Z = -3.772$; $p < 0.001$). According to Khelil et al. (2019), fear of adverse consequences may discourage internal auditors from reporting fraud or taking action in response to irregularities, while support from management and the independence of the IA function positively influence their willingness to act.

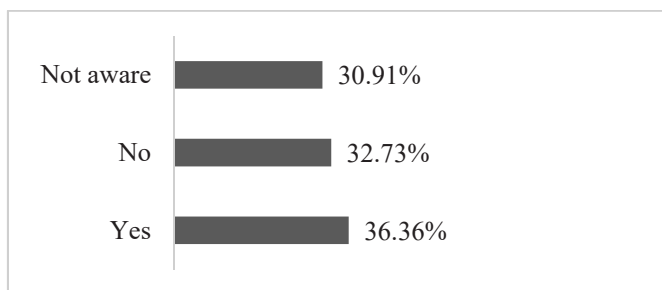
The section of the questionnaire examining the role of IA in managing greenwashing risk was completed only by respondents from organisations that report on sustainability, resulting in a subsample of 44 respondents. Respondents rated the extent to which greenwashing risk is considered within their organisations' overall risk management systems with a mean score of 2.32, while IA competencies for assessing this risk received a slightly higher rating (mean = 2.48). Regarding specific IA activities (Table 7), the highest mean score was recorded for assessing controls over sustainability information (mean = 2.25), while advisory engagement in greenwashing risk management was the least prevalent activity (mean = 1.95). However, since our study did not directly measure organisations' ESG maturity, we cannot determine whether the low level of IA engagement reflects insufficient maturity of ESG systems or the current state of IA practices in this area. A statistically significant positive correlation was found between the extent to which greenwashing risk is considered and IA competencies in this area ($\rho = 0.501$; $p < 0.001$). However, it is not possible to establish whether greater attention to this risk encourages competency development or whether more competent IA functions contribute to its greater recognition within organisations.

Table 7: Internal audit involvement in greenwashing-related activities

Activity	Mean	Std. Deviation
Assessment of greenwashing risk	2.09	1.178
Assessment of controls over the collection and processing of information underlying sustainability claims	2.25	1.184
Assessment of the consistency of published sustainability claims with internal data, policies, and processes	2.07	0.998
Advisory services on managing risks related to greenwashing	1.95	1.160

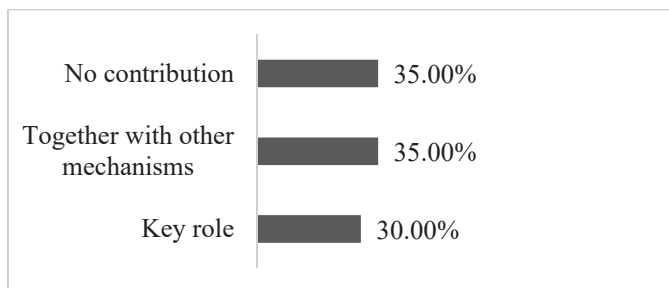
Source: Authors' presentation

Finally, internal auditors were asked about their organisations' previous experiences with fraud (Figure 3). Over the past three years, at least one significant fraud case had been detected in the organisations of 20 respondents (36.36%), while 18 respondents (32.73%) reported that no such fraud had been detected. The remaining respondents were unaware of whether any significant fraud had been detected in their organisations during the period under review.

Figure 3: Significant fraud detected during the previous three years (N=55)

Source: Authors' presentation

Among respondents who reported previous experience with significant fraud, approximately two-thirds (65%) indicated that IA had contributed to its detection, either as the key actor (30%) or alongside other detection mechanisms (35%). This finding is consistent with previous evidence of a positive association between the presence of an IA function and fraud detection (Coram, 2008). In the remaining seven cases (35%), IA did not contribute to fraud detection (Figure 4).

Figure 4: Internal audit contribution to fraud detection (N=20)

Source: Authors' presentation

Conclusion

Although IA does not have primary responsibility for preventing and detecting fraud, the professional framework incorporates the consideration of fraud risk into almost every aspect of its work (IIA, 2024). Accordingly, its contribution may vary from providing assurance on the adequacy of controls and risk management processes, through advising management and promoting an ethical culture to participating in fraud detection and, where supported by an appropriate mandate and competencies, fraud investigations. Nevertheless, the mere existence of an IA function does not guarantee its actual contribution. The extent and nature of its engagement may vary across organisations depending on its position, competencies and available resources, as well as management support and responsiveness.

Our study contributes to the literature by examining IA's role in addressing fraud and the conditions shaping this role in Serbia where the profession has a relatively short tradition compared with developed countries. The findings suggest that the potential of IA in addressing fraud has not yet been fully realised and that its role remains predominantly focused on preventive activities, while its limited engagement in other activities indicates considerable room for improvement. More frequent fraud-related training is associated with higher competency levels, while stronger competencies are associated with a greater level of engagement in fraud-related activities. A clearer mandate accompanies greater participation in assessing and improving anti-fraud controls, while adequate management responsiveness is linked to recommendations and support for developing an ethical culture. Technological resources are also associated with greater attention to fraud risk.

However, competing IA priorities and insufficiently developed technology emerged as the most prominent barriers, while internal auditors in public sector perceived pressure or adverse consequences as a significantly greater constraint. In light of these findings, management, IA functions and professional associations should focus on strengthening fraud-related competencies through continuous training, improving technological resources, clarifying IA responsibilities and ensuring adequate management support and protection from pressure.

Although the survey methodology is seen as a valid approach to examining this topic, self-perception bias among the participating internal auditors represents a potential limitation of the study. The relatively small sample and the inability to determine whether some respondents belong to the same organisation should also be considered when interpreting the findings. Future studies could further explore this topic from the perspectives of relevant stakeholders, including management, audit committees and other functions involved in addressing fraud, and combine survey data with data obtained from other sources. Further research could also provide a more detailed analysis of sector-specific differences, as well as the developing role of IA in addressing greenwashing risk.

References

- ACF. (2026). *The role of the internal auditor: Assessing and responding to fraud risk*. Anti-Fraud Collaboration.
- ACFE & Grant Thornton. (2022). *Managing Fraud Risks in an Evolving ESG Environment*. Association of Certified Fraud Examiners.
- ACFE. (2026). *Occupational Fraud 2026: A Report to the Nations*. Association of Certified Fraud Examiners.
- Allegrini, M., & D'onza, G. (2003). Internal auditing and risk assessment in large Italian companies: an empirical survey. *International Journal of Auditing*, 7(3), 191-208. <https://doi.org/10.1046/j.1099-1123.2003.00070.x>
- Anderson, U. L., Head, M. J., Ramamoorti, S., Riddle, C., Salamasick, M., & Sobel, P. J. (2017). *Internal auditing: Assurance & advisory services* (4th ed. ed.). Internal Audit Foundation.
- Asare, S. K., Davidson, R. A., & Gramling, A. A. (2008). Internal auditors' evaluation of fraud factors in planning an audit: The importance of audit committee quality and management incentives. *International Journal of Auditing*, 12(3), 181-203. <https://doi.org/10.1111/j.1099-1123.2008.00379.x>
- Asiedu, K. F., & Deffor, E. W. (2017). Fighting Corruption by Means of Effective Internal Audit Function: Evidence from the Ghanaian Public Sector. *International Journal of Auditing*, 21(1), 82-99. <https://doi.org/10.1111/ijau.12082>
- Bonrath, A., & Eulerich, M. (2024). Internal auditing's role in preventing and detecting fraud: An empirical analysis. *International Journal of Auditing*, 28(4), 615-631. <https://doi.org/10.1111/ijau.12342>
- CAQ & KRC Research. (2025). *Institutional investor survey: Research findings—Fraud*. Center for Audit Quality & KRC Research.
- Carcello, J. V., Hermanson, D. R., & Raghunandan, K. (2005). Changes in internal auditing during the time of the major US accounting scandals. *International Journal of Auditing*, 9(2), 117-127. <https://doi.org/10.1111/j.1099-1123.2005.00273.x>

- Coram, P. F. (2008). Internal audit, alternative internal audit structures and the level of misappropriation of assets fraud. *Accounting & Finance*, 48(4), 543-559. <https://doi.org/10.1111/j.1467-629X.2007.00247.x>
- COSO & ACFE. (2023). *Fraud risk management guide (2nd ed.)*. Committee of Sponsoring Organizations of the Treadway Commission.
- Deloitte Consulting AG. (2025). *2026 Internal Audit Hot Topics*. Deloitte Development LLC.
- DeZoort, F. T., & Harrison, P. D. (2008). *An evaluation of internal auditor responsibility for fraud detection*. The Institute of Internal Auditors Research Foundation.
- DeZoort, T., & Harrison, P. (2018). Understanding auditors' sense of responsibility for detecting fraud within organizations. *Journal of Business Ethics*, 149(4), 857-874.
- Drogalas, G., Pazarskis, M., Anagnostopoulou, E., & Papachristou, A. (2017). The effect of internal audit effectiveness, auditor responsibility and training in fraud detection. *Accounting and Management Information Systems*, 16(4), 434-454. <https://doi.org/10.24818/jamis.2017.04001>
- Eulerich, M. (2021). The new three lines model for structuring corporate governance – A critical discussion of similarities and differences. *Corporate Ownership & Control*, 18(2), 180-187.
- Eulerich, M., Bonrath, A., & Lopez Kasper, V. I. (2022). Internal auditor's role in ESG disclosure and assurance: An analysis of practical insights. *Corporate Ownership & Control*, 20(1), 78-86.
- Halbouni, S. S. (2015). The role of auditors in preventing, detecting, and reporting fraud: The case of the United Arab Emirates (UAE). *International Journal of Auditing*, 19(2), 117-130. <https://doi.org/10.1111/ijau.12040>
- Hazami-Ammar, S. (2019). Internal auditors' perceptions of the function's ability to investigate fraud. *Journal of Applied Accounting Research*, 20(2), 134-153. <https://doi.org/10.1108/JAAR-09-2017-0098>
- Hermanson, D. R., Smith, J. L., & Stephens, N. M. (2012). How effective are organizations' internal controls? Insights into specific internal control elements. *Current Issues in Auditing*, 6(1), A31-A50. <https://doi.org/10.2308/ciia-50146>
- IIA. (2019). *Fraud and Internal Audit: Assurance Over Fraud Controls Fundamental to Success*. The Institute of Internal Auditors.
- IIA. (2021). *Internal audit's role in ESG reporting: Independent assurance is critical to effective sustainability reporting [White paper]*. The Institute of Internal Auditors.
- IIA. (2024). *Global internal audit standards*. The Institute of Internal Auditors.

IIA. (2025). *Internal Auditing Competence Framework: Global Practice Guide*. The Institute of Internal Auditors.

Internal Audit Foundation. (2025). *Risk in Focus 2026: Global summary*. Internal Audit Foundation.

Jokipii, A., & Rautiainen, A. (2024). Quo vadis, internal auditing? A vision for internal auditing in 2030. *International Journal of Auditing*, 28(1), 170-184. <https://doi.org/10.1111/ijau.12324>

Jones, M. J., & Solomon, J. F. (2010). Social and environmental report assurance: Some interview evidence. *Accounting forum*, 34(1), 20-31. <https://doi.org/10.1016/j.accfor.2009.11.002>

Khelil, I., Akrou, O., Hussainey, K., & Noubbigh, H. (2018). Breaking the silence: An empirical analysis of the drivers of internal auditors' moral courage. *International Journal of Auditing*, 22(2), 268-284. <https://doi.org/10.1111/ijau.12119>

KPMG. (2024). *The move to mandatory reporting*. KPMG.

Ma'ayan, Y., & Carmeli, A. (2016). Internal audits as a source of ethical behavior, efficiency, and effectiveness in work units. *Journal of Business Ethics*, 137, 347–363. <https://doi.org/10.1007/s10551-015-2561-0>

Mutschmann, M., Hasso, T., & Pelster, M. (2022). Dark triad managerial personality and financial reporting manipulation. *Journal of Business Ethics*, 181(3), 763-788. <https://doi.org/10.1007/s10551-021-04959-1>

Nadirisyah, Indriani, M., & Mulyany, R. (2024). Enhancing fraud prevention and internal control: the key role of internal audit in public sector governance. *Cogent Business & Management*, 11(1), 2382389. <https://doi.org/10.1080/23311975.2024.2382389>

Novatiani, R. A., Kusumah, R. W., Yadiati, W., Abdul Halim Rachmat, R., & Arifian Rachman, A. (2024). Internal auditor competence and internal control: Improving internal audit quality to prevent fraudulent financial statements. *Cogent Business & Management*, 11(1), 2409339. <https://doi.org/10.1080/23311975.2024.2409339>

PricewaterhouseCoopers (PwC). (2023). *Global Internal Audit Study 2023: Seeing through walls to find new horizons*. PwC.

PwC. (2024). *PwC's Global Economic Crime Survey 2024 - Meeting tomorrow's challenges, embracing risk intelligently*. PricewaterhouseCoopers .

PwC. (2025). *PwC's Global Compliance Survey 2025 - Moving faster: Reinventing compliance to speed up, not trip up*. PricewaterhouseCoopers .

Rakipi, R., & D'Onza, G. (2023). The involvement of internal audit in environmental, social, and governance practices and risks: Stakeholders' salience and insights from audit committees and chief executive officers. *International Journal of Auditing*, 28(2), 522-535. <https://doi.org/10.1111/ijau.12341>

Ruiz-Blanco, S., Romero, S., & Fernandez-Feijoo, B. (2022). Green, blue or black, but washing—What company characteristics determine greenwashing? . *Environment, Development and Sustainability*, 24(3), 4024-4045.

Siano, A., Vollero, A., Conte, F., & Amabile, S. (2017). “More than words”: Expanding the taxonomy of greenwashing after the Volkswagen scandal. *Journal of Business Research*, 71, 27-37.

Transparency International. (2026). *Corruption Perceptions Index 2025*. Transparency International.

Vitalis, A., Boritz, J. E., & Simeoni, L. (2024). Enhancing CPA competencies for internal audit roles. *International Journal of Auditing*, 28(3), 458-484. <https://doi.org/10.1111/ijau.12337>

Watts, O., Elder, R., & Hyman, M. (2026). Voluntary adoption of internal audit by NASDAQ firms and its impact on internal control reporting. *International Journal of Auditing*, 30, 426–441. <https://doi.org/10.1111/ijau.70025>

Miloš Božović
University of Belgrade – Faculty of
Economics and Business
ORCID: 0000-0002-5087-1977

MARKET MANIPULATION, MOMENTUM CRASHES, AND THE LIMITS OF PRICE-VOLUME SCREENING

UDK 336.764:347.73

DOI: <https://doi.org/10.54318/czf.2026.mb.007>

Abstract

Momentum trading strategies earn persistent premia but face severe crashes during market recoveries. We investigate whether these crashes result from unwinding price manipulation. We intersect a forensic manipulation score (FMS) of abnormal turnover and illiquidity, calculated from CRSP data, with 489 SEC manipulation episodes (2000–2018). Only 14 episodes (2.9%) enter the investable momentum universe, and 94% of unmatched cases are OTC securities. In Firth penalized logit regressions, FMS loses predictive power once size is controlled, whereas Sloan operating accruals remain significant, although based on 32 events. In conditional crash regressions across 37 rebound months between 1980 and 2024, high-FMS stocks exhibit no detectable loser crash amplification (minimum detectable effect is 1.45). We therefore find no evidence that manipulation drives momentum crashes. Observable enforcement should not be treated as representative of investable equities for forensic screening, audit risk assessment, or market surveillance.

Keywords: momentum, SEC enforcement, illiquidity, operating accruals, limits to arbitrage

JEL Classification: G14, G12, K22, G18, M41

Introduction

Momentum-based trading strategies consist of buying recent winners and shorting recent losers (Jegadeesh & Titman, 1993). They earn persistent premia but suffer severe “momentum crashes” when loser stocks recover after prolonged market declines. Daniel & Moskowitz (2016) attributed momentum premia to time-varying systematic risk and limits to arbitrage. Forensic finance offers an alternative view, as extreme past returns may also reflect trade-based manipulation. If momentum portfolios contain manipulated stocks, crashes could reflect the unwinding of artificial pricing rather than a reaction to fundamental news.

In this paper, we examine the relationship between enforced manipulation and investable equities used in momentum strategies. We also explore whether a price-volume score identifies manipulation or merely proxies for scale and illiquidity. To this end, we benchmark a metric called the forensic manipulation score (FMS)

calculated from CRSP data against 489 manipulation episodes across 21,493 SEC releases from 2000 to 2018. Using Firth-penalized logit, we test whether FMS predicts enforcement beyond scale and accounting flags, specifically Sloan accruals and asset growth. Finally, we estimate conditional crash regressions across rebounds following Daniel & Moskowitz (2016).

Our central finding reveals that only 14 manipulation episodes (2.9%) survive into investable equities, defined as non-microcap stocks with a price of at least \$5, and an audit confirms that 94.0% of unmatched cases are OTC securities. Observable manipulation thus lies almost entirely outside momentum portfolios. FMS predicts enforcement alone but loses significance once size and microstructure controls are added. Sloan operating accruals remain significant while asset growth is marginal. In crash regressions, high-score stocks show no detectable amplification of losses on the loser leg. We therefore find no reliable evidence linking price manipulation to momentum crashes, and observable enforcement appears too sparse in investable equities to plausibly drive them.

This study makes three contributions. First, it documents institutional segmentation: observable enforcement rarely reaches investable equities that truly matter implementing momentum strategies. Second, it shows that price-volume heuristics add no detectable information beyond scale and microstructure, consistent with FMS proxying for illiquidity rather than capturing manipulation; on the other hand, accounting flags retain predictive power. Third, it finds no detectable amplification of momentum crashes among manipulation-prone stocks. The results are compatible with the limits-to-arbitrage explanation of momentum profits, although we do not test it directly in this paper.

1. Literature review and hypotheses

Daniel & Moskowitz (2016) show momentum crashes occur during rebounds following prolonged market declines. Distressed past losers behave like call options. When markets recover, losers achieve significant gains, inflicting heavy losses on the short leg of momentum portfolios. These losses may be partially alleviated through volatility timing strategies (Barroso & Santa-Clara, 2015).

In contrast, Allen & Gale (1992) and Aggarwal & Wu (2006) show that promoters in pump-and-dump schemes raise prices in illiquid stocks before dumping shares. Putniņš (2012) surveys how wash trades and ramping generate artificial volume and price run-ups. When schemes reverse, manipulated stocks could enter winner portfolios during run-ups and loser portfolios after collapse, offering an alternative explanation of momentum crashes. However, detecting manipulation from market data is difficult. Price-volume screens based on abnormal turnover and illiquidity (Amihud, 2002) offer an intuitive detection method but may predominantly proxy for scale and illiquidity.

Enforcement records provide the primary benchmark for manipulation, but they reflect regulatory detection choices. Manipulators typically avoid high-attention stocks of large companies with high institutional ownership. Instead, they target lightly regulated, low-priced securities where order imbalances move prices easily. This also implies authorities face lower evidentiary hurdles in documenting promotional fraud in such stocks.

Accounting signals are another detection concern, as they may blur the effect. Sloan (1996) shows high operating accruals predict lower earnings persistence and reversals. Cooper, Gulen, & Schill (2008) show rapid asset growth predicts negative abnormal returns. Following Sloan (1996) and Dechow et al. (2011), we use operating accruals and asset growth as accounting controls and test whether price-volume metrics add independent information.

We test two hypotheses.

H₁: A forensic manipulation score (FMS) built from abnormal turnover and illiquidity positively predicts subsequent SEC enforcement actions, incremental to firm size and accounting signals.

H₂: In rebound states following major declines, high-score losers show a larger rebound beta than low-score losers, which amplifies momentum crash losses.

If FMS captures manipulation, it should retain its predictive power after controlling for size, and high-score losers should exhibit larger rebound increases. In contrast, if confirmed cases lie outside investable equities and FMS adds no incremental validity beyond size, then high-score portfolios reflect illiquidity rather than manipulation.

2. Data and methodology

We retrieve market data from CRSP. We focus on common equity (codes 10/11) traded on NYSE, AMEX and NASDAQ between 1980 and 2024. NASDAQ volume uses Gao & Ritter (2010) divisors: 2.0 before Feb 2001, 1.8 in 2001, 1.6 in 2002–2003, and 1.0 post-2004. Missing delisting returns (codes 500, 520–584) are imputed at –30% for NYSE/AMEX and –55% for NASDAQ, following Shumway & Warther (1999). Screens use prior month-end prices P_{t-1} across four tiers: Tier 0 (unfiltered CRSP); Filter A (excluding NYSE size decile 1); Filter B ($P_{t-1} \geq \$1.00$); and Filter C (investable momentum: $P_{t-1} \geq \$5.00$ and market equity above the NYSE size decile 1 breakpoint; Jegadeesh & Titman, 2001).

The forensic manipulation score is an equal-weighted rank heuristic that excludes past returns:

$$FMS_{i,t} = \frac{1}{2} [\text{RankSize}(ATURN_{i,t}) + \text{RankSize}(ILLIQ_{i,t})],$$

where $ATURN_{i,t}$ is the ratio of average daily turnover of stock i from months $t - 3$ to $t - 1$ to average daily turnover from months $t - 12$ to $t - 4$. We require at least 9 months of valid volume. $ILLIQ_{i,t}$ is the Amihud (2002) measure, the trailing 12-month average of the daily ratio of absolute return to dollar volume, with at least 150 trading days. RankSize is the percentile rank within NYSE size quintiles, scaled to a unit interval, $[0, 1]$. For H_2 , we form momentum portfolios in the usual way (formation months $t - 12$ to $t - 2$, skipping month $t - 1$ to avoid short-term reversals) and split the loser decile into high- and low-FMS groups at the median. The median is computed within the loser decile each month, and portfolios are rebalanced monthly.

We index 21,493 SEC releases from 2000 to 2018, publicly available via SEC archives (sec.gov). The dataset includes 7,984 Litigation Releases, 12,543 Administrative Proceedings, and 966 Section 12(k) suspensions. We use alleged manipulation initiated in 2000–2015 to maintain a three-year lag. Standalone suspension review excluded 92.0% that were Section 13(a) reporting delinquencies. A reliability sample of 250 releases achieved 98.0% agreement. A manual audit of 50 unflagged releases found two manipulations, for an estimated miss rate of 4.0% with a 95% Wilson confidence interval of 1.1–13.5%. Applied to the 18,651 unflagged releases, this implies roughly 746 additional manipulation releases with a range of 205–2,518, so the number of undetected episodes is highly uncertain. Confirmed episodes therefore represent a lower bound on the episode count, not a direct correction to the 489 episodes or the 2.9% overlap share. Matching identifies 108 CRSP episodes in Tier 0. A manual audit of 50 unmatched cases confirms 94.0% are OTC or Pink Sheet stocks, and 6.0% reflect linkage failures. For the H_1 test, we compute FMS at month $\tau - 2$, where τ is the first month of the alleged manipulation window, so the score uses no information from $\tau - 1$. Of the 108 matched episodes, 69 have the required 9 months of pre-window history, 58 have complete microstructure data, and 32 have valid lagged accounting data (Model 3).

We use Sloan (1996) operating accruals and Cooper et al. (2008) asset growth from the Open Source Asset Pricing database (Chen & Zimmermann, 2022). We require that accounting variable observations are dated at least four months after fiscal year-end. For H_1 , we estimate Firth (1993) penalized logistic regressions on firm-year panels (Model 1: $N = 84,210$, event rate 0.082%; Model 2: $N = 81,450$, 0.071%; Model 3: $N = 62,380$, 0.051%) and report average precision (AP) with 95% bootstrap confidence intervals. For H_2 , we estimate the Daniel & Moskowitz (2016) conditional regression:

$$R_{k,t}^e = \alpha_{0k} + \alpha_{Bk}I_{B,t-1} + [\beta_{0k} + I_{B,t-1}(\beta_{Bk} + I_{U,t}\beta_{BUk})]R_{M,t}^e + \varepsilon_{k,t},$$

where $R_{k,t}^e$ is the value-weighted excess return of group k , $R_{M,t}^e$ is the market excess return (the CRSP value-weighted return minus the risk-free rate), $I_{B,t-1}$ equals 1 if the trailing 24-month market return is negative, and $I_{U,t}$ equals 1 if the month's excess market return is non-negative. We estimate it separately for k representing

high- and low-FMS losers and test the difference in β_{BUK} . We use Newey & West (1987) standard errors with 6 lags. The 1980–2024 sample has 540 months, 37 of which are rebound months, so we report standard errors, 95% confidence intervals, and the smallest β_{BUK} difference detectable at 80% power. The H_1 enforcement test uses the 2000–2018 overlap, while the H_2 crash regressions use 1980–2024.

3. Empirical results

Table 1 presents the screening funnel and overlap diagnostic. From 21,493 indexed releases (2000–2018), keyword screening flagged 2,765 candidates and manual review identified 77 non-delinquency suspensions (8.0% of 966 suspensions; 92.0% delinquent excluded), yielding a screened pool of 2,842 audited releases (13.2%). Full-text review excluded 2,248 non-manipulation releases, false hits, or out-of-window cases, leaving 594 confirmed releases consolidating into 489 unique issuer-episodes identified from releases in 2000–2018.

Panel A shows confirmed manipulation is heavily confined to unlisted venues: 381 episodes (77.9%) have no CRSP match. An audit of trading venues confirms that 94.0% of unmatched cases are verified OTC or Pink Sheets, while 6.0% reflect linkage friction. Only 108 episodes (22.1%) correspond to CRSP common stocks.

Applying standard filters quickly reduces this sample. Filter A retains 32 episodes, Filter B retains 41 episodes, and Filter C retains 14 episodes. Filter B retains more episodes than Filter A because many penny-stock targets trade between \$1.00 and \$3.00 during run-ups, yet their capitalization remains below the first NYSE decile. Adding linkage friction misses yields at most 17 Filter C survivors.

Panel B details manipulation modalities: pump-and-dump schemes comprise 68.5% of episodes (335 cases), but only five such cases survive Filter C. Order-book spoofing and layering account for 14.1% (69 cases). Six out of 14 Filter C survivors are spoofing cases, consistent with post-2010 surveillance of liquid order books following the Dodd-Frank Act. Detection bias favors visible OTC frauds. Low overlap indicates that observable enforcement rarely enters investable portfolios, though unobserved manipulation in liquid assets cannot be ruled out.

Table 2 reports penalized logistic regressions predicting SEC enforcement across CRSP calibration panels. In Model 1, unadjusted FMS yields a positive, significant coefficient, with an odds ratio of 31.28 across the full rank range. In Model 2, adding log market equity, idiosyncratic volatility (IVOL), and ATURN flips FMS's sign and eliminates its significance, while size enters negatively.¹ This suggests that price-volume anomalies add no explanatory power beyond scale and microstructure

¹ Idiosyncratic volatility (IVOL) is the annualized standard deviation of daily residuals from the three-factor model of Fama & French (1993) over the trailing month. Because ATURN enters individually in Models 2–3 while forming an FMS component, we check the variance inflation factor for collinearity and verify it remains below 2.5.

controls. Because FMS is ranked within NYSE size quintiles, its within-quintile variation still loads on scale, especially for the smallest stocks.

In Model 3, Sloan operating accruals are positive and significant, while asset growth is marginal, based on 32 events. Model 3 achieves an average precision of 0.1910 against a baseline of 0.00051, versus 0.00082 in Model 1 and 0.00071 in Model 2, which is a 375-fold gain driven by size and accruals. The equal number of events between Model 3 and Filter A survivors is coincidental: Filter A screens for size, while Model 3 reflects the availability of accounting data. H_1 is thus not supported beyond size and microstructure controls: heuristics proxy microcap scale, whereas accounting red flags retain information. As FMS lacks incremental predictive validity for enforcement beyond size, high-score stocks are more plausibly high-turnover, illiquid stocks than confirmed manipulation targets, even within the investable universe used for H_2 .

Table 1: SEC enforcement sample funnel and market overlap diagnostic

Panel A: Enforcement sample funnel and multi-tier market filters

Step / screening filter	Regulatory universe / Criterion	Sample count (N)	% Initial releases	% Confirmed episodes
Total archive indexed	SEC Master Index (2000–2018)	21,493	100.0	—
Litigation releases (LR)	Federal district court civil actions	7,984	37.1	—
Administrative proceedings (AP)	Commission administrative orders	12,543	58.4	—
Section 12(k) suspensions	Summary trading suspension orders	966	4.5	—
Keyword-screened candidate pool	Regex candidates (2,765) + Screened suspensions (77)	2,842	13.2	—
of which: Suspensions screened	Section 12(k) orders (8.0% of 966; 92.0% delinquent excluded)	77	0.36	—
Confirmed manipulation episodes	Unique issuer-episodes (releases 2000–2018; initiation 2000–2015, lower bound)	489	2.28	100.0

Unlisted OTC / Pink Sheets	Non-CRSP microcap / OTCBB (verified ~94%)	381	1.77	77.9
Tier 0: CRSP matched	CRSP common stocks (Share codes 10/11)	108	0.50	22.1
Filter A: Size screen	Exclude bottom NYSE size decile	32	0.15	6.5
Filter B: Price floor	Prior month-end price $P_{t-1} \geq \$1.00$	41	0.19	8.4
Filter C: Investable momentum	Combined: $P_{t-1} \geq$ \$5.00 AND non- microcap	14	0.07	2.9

Notes: Unit of observation is the issuer-episode. Tier 0 comprises common equity traded on NYSE, AMEX or NASDAQ. Filters A, B, and C evaluate presence in CRSP at initiation. Filter A excludes the bottom decile of NYSE size. Filter B applies $P_{t-1} \geq \$1.00$. Filter C defines investable momentum ($P_{t-1} \geq \$5.00$, market equity above the NYSE size decile 1 breakpoint).

Panel B: Distribution by manipulation modality and Filter C survival

Manipulation modality	Primary behavioral marker	Confirmed (N)	% Confirmed	Filter C survivors (N)	Filter C survival rate (%)
Pump-and- dump / Promotional hype	Microcap touting, newsletters, spam faxes	335	68.5	5	1.5
Order-book spoofing & layering	Algorithmic spoofing, marking close/open	69	14.1	6	8.7
Wash trading & matched orders	Prearranged trades, circular ownership	48	9.8	2	4.2
Suspension frenzy	Standalone 12(k) suspension for frenzy	37	7.6	1	2.7
Total Confirmed Episodes	All manipulation modalities	489	100.0	14	2.9

Notes: Double-blind coding reliability: pre-resolution kappa = 0.960; population-weighted kappa = 0.923. Filter C survival measures episodes satisfying both non-microcap size and $P_{t-1} \geq \$5.00$.

Table 2: Predictive validity of the forensic score

Independent variable	Model 1 (score only)	Model 2 (+ microstructure)	Model 3 (+ accounting red flags)
Forensic score (FMS)	3.443*** (5.68)	-1.109 (-1.17)	-1.925 (-1.51)
Odds ratio [exp(beta)]	[31.28]	[0.33]	[0.15]
Log market equity (Size)	—	-0.857*** (-10.24)	-0.859*** (-10.25)
Idiosyncratic volatility (IVOL)	—	0.082 (0.91)	0.081 (0.90)
Abnormal turnover (ATURN)	—	0.165* (1.75)	0.164* (1.74)
Sloan operating accruals	—	—	11.170*** (4.86)
Asset growth	—	—	0.845* (1.94)
Constant	-7.521*** (-18.40)	-4.218*** (-8.12)	-4.350*** (-8.21)
Firm-year observations (<i>N</i>)	84,210	81,450	62,380
Confirmed enforcement events	69	58	32
Sample prevalence (π)	0.082%	0.071%	0.051%
Pseudo- R^2	0.015	0.048	0.062
Average precision (AP)	0.00168	0.12646	0.19103
95% bootstrap CI	[0.00123, 0.00333]	[0.04893, 0.21766]	[0.07890, 0.32858]
Baseline prevalence rate	0.00082	0.00071	0.00051

Notes: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Models are estimated by Firth penalized-likelihood logistic regression to address rare events and separation. Reported z-statistics and p-values are based on the penalized likelihood (Wald statistics from the penalized information matrix). Standard errors are not clustered. Bracketed odds ratios are exp(beta) per unit change (for FMS, the full 0–1 rank range). Odds ratios per one standard deviation are reported in the text. Pseudo- R^2 is McFadden's penalized likelihood ratio index.

Table 3 reports Daniel & Moskowitz (2016) conditional crash regressions across 1980–2024. It tests whether illiquidity and turnover amplify momentum crash losses during recoveries. The sample includes 540 months: 63 bear months and 37 rebound months. For aggregate portfolios, estimates align directionally with standard patterns but are statistically insignificant: loser portfolios show higher upside rebound beta, while the momentum portfolio (WML) shows an insignificant negative interaction. When stratified by FMS, the difference portfolio $\Delta\text{Loser} = \text{Loser}_{\text{High-FMS}} - \text{Loser}_{\text{Low-FMS}}$ yields $\beta_{B,U} = 0.324$, while $\beta_{B,U} = -0.330$ for ΔWML . However, neither difference coefficient is statistically significant. Reporting standard errors and confidence

intervals highlights the test's power constraints. For ΔLoser , the 95% confidence interval is relatively wide, and the Minimum Detectable Effect at 80% power is 1.45. With only 37 rebound months, standard errors are considerable, and the test lacks power to detect moderate crash amplification. Insignificance reflects low power rather than definitive proof of absence.

The results yield a consistent verdict. First, Table 1 shows observable regulatory manipulation is segregated in OTC microcaps, with negligible presence in investable momentum. Second, Table 2 shows price-volume heuristics fail to predict enforcement beyond size and microstructure controls. Third, Table 3 finds that while the positive point estimate of 0.324 for ΔLoser is economically intuitive and consistent in sign with H_2 , it is insignificant. We therefore find no statistically reliable evidence that momentum crashes are driven by trade-based price manipulation. Rebound drawdowns in investable equities are more plausibly governed by limits to arbitrage, time-varying market risk premiums, and operating distress, though we do not test these channels.

Our study has several methodological limitations that may influence its conclusions. First, regulatory archives capture observable enforcement rather than *all* misconduct, so detection bias favors visible OTC schemes over manipulation in highly liquid order books. If missed cases disproportionately appear on major exchanges or involve larger, more liquid firms, the true overlap between manipulation and the investable momentum universe would be higher than the 2.9% we report. If they are mostly OTC, it would be lower. The 4.0% miss rate rests on 2 of 50 audited unflagged releases. Taken at face value, it implies that our keyword screen recovered only about 44% of manipulation releases. The overlap should therefore be read as conditional on SEC detection and our keyword screen, and the true number of manipulation episodes may be materially larger than 489. Second, accounting tests in Model 3 rely on 32 confirmed events with standardized accounting coverage, which limits precision. Third, crash regressions are constrained by 37 rebound months across 1980–2024, yielding wide confidence intervals and an MDE of 1.45 that cannot rule out moderate-to-large effects. Finally, our OTC concentration estimate relies on a hand audit of 50 unmatched cases, carrying modest sampling uncertainty.

Table 3: Daniel & Moskowitz market-state rebound regressions across forensic manipulation strata (1980–2024)

Portfolio Specification	α_0 (%)	β_0	β_B	$\beta_{B,U}$ (SE)	95% Conf. Interval	t - stat	p -val	MDE	R ²
Winner (All)	0.20	1.113	– 0.056	–0.282 (0.158)	[–0.59, 0.03]	– 1.78	0.075	0.44	0.852
Loser (All)	– 0.54	1.130	0.382	0.263 (0.386)	[–0.50, 1.02]	0.68	0.497	1.08	0.770
WML (All)	0.74	– 0.017	– 0.438	–0.544 (0.479)	[–1.48, 0.39]	– 1.14	0.255	1.34	0.147
Winner (Low FMS)	0.15	0.854	– 0.136	–0.123 (0.171)	[–0.46, 0.21]	– 0.72	0.473	0.48	0.677
Winner (High FMS)	0.18	1.424	– 0.097	–0.129 (0.359)	[–0.83, 0.58]	– 0.36	0.720	1.01	0.681
Loser (Low FMS)	– 0.00	0.736	0.321	0.202 (0.341)	[–0.47, 0.87]	0.59	0.554	0.96	0.470
Loser (High FMS)	– 1.15	1.466	0.378	0.526 (0.529)	[–0.51, 1.56]	0.99	0.320	1.48	0.709
Δ Loser (High – Low)	– 1.14	0.729	0.057	0.324 (0.516)	[–0.69, 1.34]	0.63	0.530	1.45	0.268
WML (Low FMS)	0.16	0.118	– 0.457	–0.325 (0.432)	[–1.17, 0.52]	– 0.75	0.452	1.21	0.061
WML (High FMS)	1.32	– 0.041	– 0.476	–0.655 (0.552)	[–1.74, 0.43]	– 1.19	0.236	1.55	0.157
Δ WML (High – Low)	1.17	– 0.159	– 0.018	–0.330 (0.343)	[–1.00, 0.34]	– 0.96	0.336	0.96	0.051

Notes: Newey-West SEs with 6 lags. $N = 540$ months (1980–2024), including 37 rebound months. $\beta_{B,U}$ measures asymmetric upside market exposure during recoveries following bear markets. Minimum Detectable Effect (MDE) is computed at 80% power. Winner and Loser portfolios are formed as in Daniel & Moskowitz (2016). WML is the momentum portfolio (Winner – Loser).

4. Forensic and practical implications

For institutional asset managers, our results suggest that price-volume manipulation screens are unlikely to hedge momentum crashes reliably. Because confirmed schemes in regulatory records are concentrated in unlisted OTC venues, price-volume heuristics used to filter institutional portfolios do not insulate portfolios from rebound drawdowns. Momentum left-tail risk is more plausibly governed by distressed firm leverage and limits to arbitrage than detectable manipulation, though we do not test this. For external auditors under ISA 240 and PCAOB AS 2401, our

findings point to quantitative risk indicators. Our evidence does not show that price-volume spikes in exchange-listed stocks indicate outright manipulation, and the small number of matched episodes limits what we can conclude. We conjecture that their co-occurrence with elevated operating accruals and rapid asset expansion may signal heightened earnings management risk. Extreme turnover may reflect speculative scrutiny that incentivizes aggressive accounting, warranting substantive testing of revenue recognition and accruals.

For surveillance authorities like the SEC or FINRA, our analysis shows the operational boundaries of price-volume surveillance. While automated turnover alerts flag microcap schemes, their lack of incremental validity beyond scale means applying them to exchange-listed markets is unlikely to improve detection. In exchange-listed equities, surveillance is more likely directed at order-book abuses, such as spoofing and layering, and financial restatements, as our six Filter C spoofing survivors show. However, because archives reflect detected cases, authorities must recognize that low overlap does not preclude undetected manipulation in liquid assets.

Conclusion

This paper evaluates whether momentum crashes are driven by trade-based capital market manipulation. By combining CRSP data between 1980 and 2024 with SEC enforcement actions, we empirically assess the intersection of trade-based manipulation, regulatory enforcement, and momentum crash dynamics.

Our results provide no statistically reliable evidence that momentum crashes are driven by trade-based price manipulation. Among enforced cases, actions are concentrated in unlisted OTC penny stocks: only 2.9% of confirmed manipulation episodes survive into the standard investable momentum universe, with an audit confirming 94.0% of unmatched cases are verified OTC venues.

The forensic price-volume score adds no detectable predictive power for regulatory enforcement once size and microstructure controls are included, suggesting that price-volume anomalies in regulatory targets may primarily reflect microcap illiquidity. In contrast, Sloan operating accruals remain significant, while asset growth is marginal, though both rest on 32 events.

Finally, in conditional Daniel & Moskowitz (2016) crash regressions, stratifying portfolios by price-volume heuristics produces statistically insignificant differences in crash beta. Because historical rebound regimes span only 37 months, the resulting 95% confidence interval is wide, with an MDE of 1.45, suggesting that the insignificance reflects the power constraints of rare recoveries rather than proof of absence.

References

- Aggarwal, R. K., & Wu, G. (2006). Stock market manipulations. *The Journal of Business*, 79(4), 1915–1954. <https://doi.org/10.1086/503652>
- Allen, F., & Gale, D. (1992). Stock-price manipulation. *The Review of Financial Studies*, 5(3), 503–529. <https://doi.org/10.1093/rfs/5.3.503>
- 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](https://doi.org/10.1016/S1386-4181(01)00024-6)
- Barroso, P., & Santa-Clara, P. (2015). Momentum has its moments. *Journal of Financial Economics*, 116(1), 111–120. <https://doi.org/10.1016/j.jfineco.2014.11.010>
- Chen, A. Y., & Zimmermann, T. (2022). Open source cross-sectional asset pricing. *Critical Finance Review*, 11(2), 207–264. <https://doi.org/10.1561/104.00000112>
- Cooper, M. J., Gulen, H., & Schill, M. J. (2008). Asset growth and the cross-section of stock returns. *The Journal of Finance*, 63(4), 1609–1651. <https://doi.org/10.1111/j.1540-6261.2008.01370.x>
- Daniel, K., & Moskowitz, T. J. (2016). Momentum crashes. *Journal of Financial Economics*, 122(2), 221–247. <https://doi.org/10.1016/j.jfineco.2015.12.002>
- Dechow, P. M., Ge, W., Larson, C. R., & Sloan, R. G. (2011). Predicting material accounting misstatements. *Contemporary Accounting Research*, 28(1), 17–82. <https://doi.org/10.1111/j.1911-3846.2010.01041.x>
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56. [https://doi.org/10.1016/0304-405X\(93\)90023-5](https://doi.org/10.1016/0304-405X(93)90023-5)
- Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1), 27–38. <https://doi.org/10.1093/biomet/80.1.27>
- Gao, C. X., & Ritter, J. R. (2010). The marketing of seasoned equity offerings. *Journal of Financial Economics*, 97(1), 33–52. <https://doi.org/10.1016/j.jfineco.2010.03.007>
- Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, 48(1), 65–91. <https://doi.org/10.1111/j.1540-6261.1993.tb04702.x>
- Jegadeesh, N., & Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *The Journal of Finance*, 56(2), 699–720. <https://doi.org/10.1111/0022-1082.00342>
- King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, 9(2), 137–163. <https://doi.org/10.1093/oxfordjournals.pan.a004868>

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3), 703–708. <https://doi.org/10.2307/1913610>

Putniņš, T. J. (2012). Market manipulation: A survey. *Journal of Economic Surveys*, 26(5), 952–967. <https://doi.org/10.1111/j.1467-6419.2011.00692.x>

Shumway, T. and Warther, V. A. (1999). The delisting bias in CRSP's Nasdaq data and its implications for the size effect. *The Journal of Finance*, 54(6), 2361–2379. <https://doi.org/10.1111/0022-1082.00192>

Sloan, R. G. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *The Accounting Review*, 71(3), 289–315. <https://doi.org/10.2308/TAR-9608042309>

Diana Dumitrescu
West University of Timisoara, Romania
ORCID: 0009-0009-0256-0291

Nicolae Bobițan
West University of Timisoara, Romania
ORCID: 0009-0008-1295-4262

Adriana Florina Popa
Bucharest Academy of Economic
Studies, Romania)
ORCID: 0000-0002-9995-3893

ARTIFICIAL INTELLIGENCE, FORENSIC ACCOUNTING, AND TAX-CRIME INVESTIGATIONS: CROSS-COUNTRY EVIDENCE

UDK 657.6:004.8:336.2

DOI <https://doi.org/10.54318/czf.2026.dd.008>

Abstract

Tax administrations have rapidly expanded their use of artificial intelligence (AI), particularly for risk assessment and the detection of tax evasion and fraud. This technological transformation has generated expectations that AI-supported tax administration should strengthen tax-crime enforcement and increase the identification of complex cases requiring forensic-accounting expertise. However, existing research remains focused predominantly on technology adoption, algorithmic performance, and individual administrative applications, while systematic cross-country evidence connecting AI implementation with observable forensic-investigative activity remains limited.

This study examines whether AI implementation in national tax administrations is associated with greater forensic accounting investigative activity across European Union (EU) Member States. Variable analyzed are revenue-fraud investigations per 100,000 inhabitants, revenue-fraud cases as a share of all active EPPO investigations, and revenue-fraud damage relative to VAT and customs revenue, based on an exploratory cross-sectional design combining descriptive statistics, percentile rankings, Welch tests, Mann-Whitney tests, and regression analysis.

AI-implementing Member States exhibit higher values across all three indicators analyzed. The most consistent evidence concerns the share of revenue-fraud cases within the EPPO caseload, although the non-parametric evidence is significant only at the 10% level. Overall, the findings are consistent with a complementary relationship in which AI supports risk identification and case prioritization, while forensic accounting expertise remains necessary to reconstruct transactions, quantify losses, and develop evidentially sustainable cases.

Key words: artificial intelligence, forensic accounting, tax administration, tax fraud detection, tax crime investigations, cross-country analysis

JEL Classification: M41, M42, H26, K42

Introduction

Tax administrations have become increasingly data-intensive institutions. Electronic tax returns, digital invoicing, third-party reporting, interconnected public databases, automated compliance systems, and large volumes of transactional information provide authorities with data that could not realistically be analyzed through conventional manual procedures alone. At the same time, tax fraud has become increasingly complex, frequently involving multiple entities, cross-border transactions, fictitious invoices, concealed ownership structures, circular transaction chains, and arrangements intended to obscure the economic substance of transactions.

Artificial intelligence has consequently emerged as an important component of the digital transformation of tax administration. AI-based systems allow authorities to process large volumes of structured and unstructured information, identify patterns that may not be visible through conventional analytical approaches, detect relationships among taxpayers and transactions, and prioritize cases according to estimated compliance or fraud risk.

The OECD reports a substantial expansion in AI implementation across tax administrations. As result, in 2023, 69.0% of the administrations covered by *Tax Administration 2025* reported that artificial intelligence, including machine learning, had been implemented and was being used, compared with 29.8% in 2018 and 8.6% in 2016. A further 24.1% reported that AI was under implementation for future use (OECD, 2025a). AI has become involved especially in compliance-risk management: 64.1% of surveyed administrations reported using AI for risk assessment, while 74.4% reported using it for detecting tax evasion and fraud (OECD, 2025a).

These developments create an expectation that AI adoption should be reflected not only in technological capability but also in others enforcement activities. More effective anomaly detection and risk selection should, in principle, increase the probability of identifying potentially fraudulent transactions or improve the quality of cases selected for further examination.

However, detection represents only the initial stage of tax-crime enforcement. A suspicious transaction, taxpayer, or transaction network identified through an algorithm does not constitute evidence of fraud. The underlying economic conduct must be reconstructed, financial relationships must be established, flows of funds must be traced, accounting and tax records must be reconciled, fiscal losses must be quantified, and the findings must ultimately be supported by sufficiently reliable documentary and financial evidence.

These activities closely correspond to the investigative and evidentiary functions of forensic accounting. Forensic accounting combines accounting expertise, financial analysis, investigative procedures, professional skepticism, and evidentiary reasoning in the examination of suspected fraud and other forms of financial

misconduct. Within tax-crime investigations, forensic-accounting functions may include examination of accounting records, reconstruction of transaction chains, tracing of payments, identification of related entities or beneficial interests, reconciliation of accounting and tax information, quantification of fiscal damage, and preparation of financial evidence for administrative or criminal proceedings (Lyeonov et al., 2024; Meneses Cerón et al., 2026; Qatawneh, 2024). These functions are already appropriately identified in the draft.

AI and forensic accounting should therefore not be understood as competing approaches to tax-fraud enforcement, operating at different but interconnected stages of the investigative process. AI primarily supports screening, anomaly detection, network identification, risk assessment, and case prioritization and forensic-accounting expertise becomes particularly relevant once a suspicious case requires transaction reconstruction, financial interpretation, loss quantification, and evidentiary substantiation.

The resulting enforcement sequence can be conceptualized as:

AI-enabled data analysis → risk identification → suspicious-case selection

Forensic-financial investigation → transaction reconstruction → damage quantification

Evidence corroboration

Prosecutorial action

Existing literature demonstrates that AI is increasingly used for tax-risk assessment and fraud detection (OECD, 2025a), that advanced technologies are becoming increasingly relevant to forensic-accounting analysis (Lyeonov et al., 2024; Meneses Cerón et al., 2026), and that enforcement outcomes are strongly influenced by institutional capacity (Adan et al., 2023; Milosavljević et al., 2024). What remains insufficiently examined is whether national tax administrations implementing AI exhibit greater downstream investigative activity in cases requiring substantial forensic-financial analysis. The present study addresses this gap by linking AI implementation to fraud investigation intensity and interpreting these investigations as an activity based variable for forensic accounting enforcement.

Direct comparable statistics identifying the number of formally designated forensic accountants participating in tax-crime investigations are not publicly available. As result, our study employs an activity-based variable to measure forensic accounting activity. EPPO revenue-fraud investigations are treated as forensic-accounting-intensive investigative activity because complex VAT and customs fraud cases

typically require substantial financial reconstruction, accounting analysis, tracing of financial flows, and quantification of financial damage.

The OECD (2025a) notes that tax administrations possess information that may be critical to criminal tax investigations, but responsibility for conducting such investigations differs substantially across jurisdictions. Some tax administrations conduct investigations directly, others operate under the authority of police or prosecutors, while in other countries tax-crime investigations are undertaken by specialized institutions outside the tax administration.

The EPPO framework reduces part of this institutional heterogeneity. Although national detection and referral mechanisms remain different, EPPO applies a common prosecutorial framework to offences affecting the financial interests of the European Union, including complex VAT and customs fraud. 24 EU Member States participated in EPPO in 2025, while Denmark, Hungary, and Ireland did not (EPPO, 2026). The scale of revenue fraud within EPPO's caseload highlights the significance of financial investigation, in 2025, EPPO reported 3,602 active investigations involving estimated damage of EUR 67.27 billion. Of these, 981 involved VAT and customs revenue fraud and represented EUR 45.01 billion in estimated damage, more than two-thirds of the total estimated damage under investigation (EPPO, 2026).

Against this background, the objective of this study is to examine whether AI implementation in national tax administrations is associated with greater forensic-accounting-intensive investigative activity in tax-crime enforcement across EU Member States, using EPPO revenue-fraud investigations as an activity-based proxy for forensic-accounting involvement.

The corresponding research question is: AI implementation in national tax administrations is associated with greater forensic-accounting-intensive revenue-fraud investigation activity?

The study contributes to the literature in three respects. First, it moves beyond technology-adoption indicators and algorithmic performance to examine observable downstream investigative outcomes. Second, it brings forensic accounting directly into the empirical framework through a functional, activity-based proxy. Third, it conceptualizes AI and forensic-accounting expertise as complementary capabilities within the tax-crime enforcement process.

1. AI-Enabled Tax Administration and Forensic

1.1. Artificial intelligence in tax administration

Digitalization has transformed tax administration from predominantly document-based systems toward increasingly automated and data-driven models of compliance management. Contemporary tax authorities draw on electronic returns, electronic invoicing, third-party reporting, customs information, interconnected public

databases, and other administrative sources. The resulting increase in both data volume and complexity has created strong incentives to employ advanced analytics and artificial intelligence.

The OECD's comparative evidence documents the rapid diffusion of these technologies. AI is increasingly used in risk assessment, detection of tax evasion and fraud, automated decision support, and taxpayer services (OECD, 2025a, 2025b, 2025c). The prominence of fraud detection among these applications creates a direct theoretical connection between AI implementation and subsequent investigative activity.

The academic literature broadly confirms the potential benefits of this technological transformation. Belahouaoui and Attak (2024), in a systematic review of digital taxation and Tax Administration 3.0, identify AI as an important mechanism for improving compliance management and administrative efficiency, while emphasizing significant variation in technological maturity and institutional readiness across countries. Zaim and Sahbani (2026) similarly conclude that AI can strengthen tax-compliance monitoring and tax-evasion mitigation but stress that effectiveness depends on data quality, organizational readiness, institutional capacity, and governance. AI adoption should therefore not be treated as equivalent to effective implementation. Anjarwi (2026) identifies AI-enabled audit automation, advanced analytics, and data-driven risk identification as important components of contemporary tax-audit transformation. Judijanto and Nurhayati (2026) likewise identify potential improvements in fraud detection and administrative performance but emphasize constraints associated with infrastructure, fragmented data, governance, and regulation. More direct empirical evidence is provided by Battaglini et al. (2025), who demonstrate using Italian administrative data that machine-learning methods can improve tax-audit targeting and the identification of tax evasion.

In conclusion, the literature demonstrates substantial potential for AI to improve tax-risk assessment, fraud detection, and case selection. However, it does not establish whether tax administrations that implement AI systematically generate greater downstream criminal-investigative activity. This distinction is particularly important from a forensic-accounting perspective, because algorithmic detection, case identification, forensic investigation, evidentiary development, and prosecution represent analytically distinct stages of the enforcement process. AI may strengthen the identification of suspicious cases, but their progression toward criminal investigation depends on the capacity to reconstruct transactions, quantify losses, corroborate evidence, and satisfy prosecutorial requirements.

1.2. Forensic accounting, tax-fraud investigation, and AI capability

Forensic accounting combines accounting, taxation, legal expertise with investigative procedures and evidentiary analysis in the examination of suspected financial transactions, being involved in cases of financial statement fraud, asset

misappropriation, corruption, money laundering, tax fraud, and other forms of economic crime.

Recent literature increasingly examines the interaction between forensic accounting and advanced analytical technologies. Lyeonov et al. (2024) document substantial growth in research applying AI and machine learning to illegal financial operations, including money laundering, corruption, and fraud. Thakkar et al. (2025) similarly identify the rapid expansion of AI-based fraud-detection research. Meneses Cerón et al. (2026) identify supervised and unsupervised machine learning, anomaly detection, natural-language processing, and related techniques as increasingly important tools within forensic-accounting research and practice. Qatawneh (2024) highlights the relevance of AI in auditing and fraud detection within accounting information systems, while Adelakun et al. (2024) discuss AI-supported techniques for improving accounting fraud detection.

Nevertheless, forensic accounting extends beyond anomaly classification. A high-risk score, unusual invoice network, or anomalous financial transaction represents an investigative lead rather than proof of fraud. Establishing fraud requires interpretation of economic substance, reconstruction of the relevant transactions, corroboration of financial and documentary evidence, and quantification of the resulting economic or fiscal damage. For tax fraud investigations, these activities may require the integration of accounting ledgers, tax returns, invoices, contracts, banking information, ownership structures, payment data, transport documentation, and third-party records.

The relationship between AI and forensic accounting is most appropriately conceptualized in terms of complementarity. AI has comparative advantages in processing large datasets, identifying hidden relationships, recognizing anomalies, analyzing networks, and prioritizing risk. Forensic-accounting expertise has comparative advantages in reconstructing economic events, interpreting financial information, quantifying losses, evaluating alternative explanations, corroborating evidence, and communicating findings within legal and investigative settings. AI may therefore increase the flow of cases reaching the forensic investigative stage, while forensic accounting capacity determines whether those cases can be developed into evidentially sustainable investigations.

Tax-crime enforcement depends on institutional conditions in addition to technological capability, relating to staffing, investigative authority, information access, technological infrastructure, legal frameworks, prosecutorial cooperation, and wider governance quality.

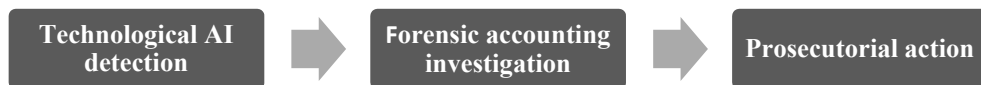
Adan et al. (2023) show that improvements in tax-administration capacity are associated with stronger revenue outcomes, while Milosavljević et al. (2024) identify a relationship between tax-administration performance and tax-evasion outcomes in European administrations. De Simone et al. (2023) similarly demonstrate that enforcement effects vary across institutional environments and taxpayer groups.

AI adoption is therefore unlikely to be randomly distributed. Administrations with greater institutional capacity may simultaneously be more capable of implementing AI and of developing complex tax-fraud investigations. The supplementary empirical specifications consequently control Government Effectiveness, Rule of Law, and Control of Corruption (Kaufmann & Kraay, 2024; World Bank, 2025b).

In conclusion, the literature suggests that AI capability and forensic-accounting expertise are complementary rather than substitutive in tax-fraud enforcement. AI strengthens anomaly detection, risk assessment, and case prioritisation, while forensic accounting supports transaction reconstruction, loss quantification, and evidentiary substantiation. Because the transition from an algorithmic signal to a prosecutable case depends on professional and institutional capacity, the relevant question is whether AI-enabled detection is associated with greater investigative activity requiring substantial forensic-accounting input.

2. Empirical Approach

The conceptual framework distinguishes between three successive stages of tax-crime enforcement:



AI is expected to affect primarily the first stage, while forensic-accounting expertise becomes particularly important in the second. At the technological-detection stage, AI can process large volumes of tax and transactional data, identify anomalies, detect hidden relationships among taxpayers and entities, and support risk scoring. Two mechanisms are particularly relevant. First, AI may improve the discovery of suspicious patterns that would otherwise remain undetected. Second, it may improve the allocation of scarce enforcement resources by directing attention toward cases with higher estimated fraud risk or financial materiality.

The empirical test is designed to assess whether AI directly causes forensic-accounting work, nor whether every EPPO investigation formally involves a professional with certification as forensic accountant, testing whether countries whose tax administrations implement AI display greater downstream revenue-fraud investigative activity, that is plausibly forensic-accounting-intensive.

The empirical relationship examined in this paper therefore concerns AI implementation and downstream revenue-fraud investigative activity that is expected to require substantial forensic-accounting functions.

Accordingly, the following hypothesis is tested:

H1. EU Member States whose tax administrations implement artificial intelligence exhibit greater forensic-accounting-intensive revenue-fraud investigation activity than Member States whose tax administrations do not implement AI.

The hypothesis is examined through three complementary dimensions: investigation volume relative to population, the concentration of revenue fraud within the EPPO caseload, and the financial materiality of the investigated revenue fraud.

2.1. Data and Variables

The initial country frame comprises the 27 Member States of the European Union.

AI implementation is identified from country level by ISORA information reported in *Tax Administration 2025* (OECD, 2025a). Under the classification used in this study, 24 EU tax administrations have implemented and use AI, including machine learning, while Croatia, Cyprus, and Czechia are classified as „non applicable”. The resulting EU-27 AI implementation rate is 88,9% higher than the 69.0% implementation rate reported for the broader group of administrations included in the OECD analysis.

Revenue-fraud outcomes are for estimation of forensic accounting activity was obtained from the *EPPO Annual Report 2025*. In 2025, 24 EU Member States participated in EPPO. Denmark, Hungary, and Ireland did not and therefore have no EPPO outcomes for inclusion in the analysis.

Consequently, Croatia, Cyprus, and Czechia are all EPPO participants and exclusion from the dependent-variable sample is unrelated to AI status. The final sample comprises 21 AI-implementing and three non-implementing Member States.

Table 1 presents variables definition, sources and formulas.

Table 1: Variables Definition and Sources

Variable	Definition	Source	Formula
AI implementation	Reported by ISORA information	OECD/ISORA	$1 = \text{implemented};$ $0 = \text{not implemented}$
Revenue-Fraud Investigation Rate (RFIR)	Number of active revenue-fraud investigations per 100,000 inhabitants	EPPO World Bank, (2025a)	$\frac{\text{RevenueFraudCases}}{\text{Population}}$
Revenue-Fraud Investigation Share (RFIS)	Proportion of each country's total active EPPO caseload represented by VAT and customs revenue fraud.	EPPO	$\frac{\text{VATRevenueFraudCases}}{\text{AllEPPOCases}}$

Variable	Definition	Source	Formula
Revenue-Fraud Financial Materiality (RFFM)	Financial magnitude of the revenue fraud under investigation relative to the corresponding fiscal base	EPPO, Eurostat (2026)	$\frac{VATRevenueFraud}{VATRevenue}$
Government Effectiveness (GEff)	Perceived quality of government credibility	World Bank	WGI estimate
Rule of Law (RoL)	Confidence in and adherence to legal rules	World Bank	WGI estimate
Control of Corruption (CC)	Extent to which public power is exercised for corruption control	World Bank	WGI estimate

Source: Authors' projection

The study therefore uses three activity-based indicators of revenue-fraud investigation that plausibly require substantial forensic-accounting functions: RFIR, RFIS and RFFM.

The analysis consequently combines complementary methods. First, means and medians are calculated separately for AI implementers and non-implementers for each dependent measure. Second, Welch's unequal-variance t-test is reported because it does not assume equal group variances. Third, the Mann-Whitney U test provides a rank-based non-parametric comparison. Particular weight is placed on this test because it does not require normality or equality of variances; nevertheless, its statistical power is necessarily limited by the presence of only three observations in the non-AI group. Fourth, the individual rank and percentile positions of Croatia, Cyprus, and Czechia are examined within the full 24 country distribution. This provides a useful descriptive assessment of whether non-implementing countries are concentrated in the upper or lower part of the outcome distribution.

Finally, supplementary OLS models are estimated:

$$Y_i = \beta_0 + \beta_1 AIImplementation_i + \beta_2 GovernmentEffectiveness_i + \beta_3 RuleOfLaw_i + \beta_4 ControlOfCorruption_i + \varepsilon_i$$

where Y_i is replaced successively by RFIR, RFIS, and RFFM.

HC1 heteroskedasticity-robust standard errors are used.

Given $N = 24$, the regressions are interpreted as supplementary specifications rather than as the principal source of inference.

Most importantly, the empirical design is associational. Neither group comparisons nor cross-sectional regressions establish a causal effect of AI implementation.

2.2. Descriptive Statistics

Table 2 presents the descriptive statistics for the main variables used in the analysis. The variables capture three complementary dimensions of forensic-accounting-intensive revenue-fraud activity: the volume of active investigations relative to population, the share of revenue-fraud cases within the total EPPO caseload, and the financial materiality of investigated fraud relative to VAT and customs revenue. The table also reports AI implementation status and the institutional variables Government Effectiveness, Rule of Law, and Control of Corruption. Considerable cross-country variation is expected across the investigative indicators, reflecting differences in country size, fraud exposure, prosecutorial activity, institutional capacity, and the scale of individual cases. These descriptive statistics therefore provide an initial overview of the distribution of the variables and support the subsequent comparison between AI-implementing and non-implementing Member States.

Table 2: Country-level Descriptive Statistics

	<i>AI</i>	<i>RFIR</i>	<i>RFIS %</i>	<i>RFFM %</i>	<i>GEff</i>	<i>RoL</i>	<i>CC</i>
Mean	0,87	0,35	27,91	3,452	6,68	7,24	5,399
Standard Error	0,06	0,06	4,115	0,905	1,52	1,55	1,574
Median	1	0,25	24,6	1,855	1,02	1,13	0,780
Standard Deviation	0,33	0,33	20,16	4,436	7,47	7,59	7,712
Sample Variance	0,11	0,11	406,4	19,68	55,8	57,6	59,48
Kurtosis	4,21	4,47	0,717	5,804	-1,37	-1,73	-0,675
Skewness	-2,4	2,05	1,004	2,354	0,60	0,39	1,096
Range	1	1,39	78,45	18,23	21,3	20,2	20,27
Minimum	0	0,04	0,45	0,21	0,10	-0,03	-0,15
Maximum	1	1,44	78,9	18,44	21,4	20,2	20,12
Sum	21	8,43	670,0	82,86	160	174	129,5
Count	24	24	24	24	24	24	24

Source: Authors' calculations

3. Results

Table 3: AI implementation and investigative outcomes

Variables	AI mean/median	Non-AI mean/median	Welch test	Mann- Whitney	OLS
RFIR	0,375/0,269	0,187/0,181	t=2,348 p=0,028	U=41,0 p=0,452	b=0,215 p=0,039 R ² =0,063
RFIS	32,75/31,03	13,25/14,29	t=3,619 p=0,004	U=52,5 p=0,074	b=15,54 p=0,007 R ² =0,416
RFFM	3,815/2,014	0,908/1,141	t=2,770 p=0,011	U=49,0 p=0,145	b=3,839 p=0,021 R ² =0,205

Source: Authors' calculations

Table 3 shows that AI-implementing countries exhibit higher means and medians for all three indicators.

For RFIR, the mean revenue-fraud investigation rate is 0,375 cases among AI implementers compared with 0,187 among non-implementers. The Welch test indicates a statistically significant difference in means ($p = 0,028$), while the OLS coefficient is positive and statistically significant ($b = 0,215$, $p = 0,039$). The Mann-Whitney test, however, does not indicate a corresponding systematic rank difference ($p = 0,452$). The divergence suggests that the mean difference is influenced by the magnitude of particular observations within the AI group rather than by a uniform displacement of the distribution.

According to RFIS results revenue-fraud cases represent, on average, 32,75% of active EPPO investigations in AI-implementing countries compared with 13,25% among the three non-implementers. The Welch test is statistically significant ($p = 0,004$), and the OLS coefficient indicates a difference of 15,54 percentage points ($p = 0,007$). The Mann-Whitney result is $p = 0,074$, providing marginal evidence at the 10% level, but not meeting a 5% threshold.

The financial-materiality indicator, RFFM, also suggest a large difference in group means. Estimated revenue-fraud damage represents 3,815% of the relevant revenue base among AI-implementing countries compared with 0,908% among non-implementers. Both the Welch comparison ($p = 0,011$) and OLS coefficient ($p = 0,021$) are statistically significant, whereas the Mann-Whitney result is not ($p = 0,145$).

Taken together, the results are consistent across all three measures but differ in distributional robustness. RFIS provides the most internally consistent evidence. RFIR and RFFM appear more sensitive to individual observations.

Because the comparison group consists of only three countries, their individual position within the complete distribution provides useful supplementary information.

Table 4: Rank of non-AI countries within the 24-country sample

Country	RFIR rank/percentiles	RFIS rank/percentiles	RFFM rank/percentiles
Croatia	15/41,7	21/16,7	17/33,3
Cyprus	18/29,2	20/22,9	22/22,5
Czechia	14/45,8	18/29,2	16/37,5

Source: Authors' calculations

All three non-implementing countries are in the lower half of the sample distribution for every outcome (Table 4). This is particularly evident for RFIS: Croatia ranks 21st, Cyprus 20th, and Czechia 18th among the 24 countries. The ranking evidence is therefore consistent with the group comparisons but should not be interpreted as independent confirmation of an AI effect. It describes the position of the three observed non-implementers within this particular sample.

H1 predicts greater forensic-accounting-intensive revenue-fraud investigation activity among AI-implementing Member States. The evidence is consistent with H1 and AI-implementing countries exhibit higher means and medians across RFIR, RFIS, and RFFM, and the OLS coefficients are positive for all three measures.

However, inferential support is not uniform. As result, parametric comparisons are statistically significant for all three indicators, whereas rank-based evidence is considerably weaker. The Mann-Whitney result provides only marginal support for RFIS at the 10% level and does not support RFIR or RFFM. Accordingly, the evidence is consistent with H1 but does not constitute strong or uniform confirmation, and the findings should be interpreted as an exploratory, hypothesis-consistent pattern rather than as confirmatory evidence of a general or causal AI effect.

4. Discussion

The empirical results show a consistent association between AI implementation and revenue-fraud investigative activity, but the strength of the evidence differs across outcome measures. The most significant result concerns RFIS, the proportion of each country's active EPPO caseload attributable to revenue fraud. Unlike the per-capita investigation rate, RFIS is a within-country compositional measure, but differences remain attributable to referral practices, institutional capacity, or fraud exposure.

The finding that RFIS is higher among AI-implementing countries is therefore compatible with a mechanism in which AI-supported risk identification is associated with a different composition of cases reaching the prosecutorial stage. If AI improves

the identification of complex VAT and customs fraud, such cases may occupy a larger share of prosecutorial activity, but interpretation remains associational. The country-level data do not reveal whether any individual EPPO investigation originated from an AI-generated risk signal and therefore cannot establish the case-level mechanism linking AI use to prosecution.

RFIR and RFFM provide weaker distributional evidence, and both are significant under the Welch and OLS specifications but not under the Mann-Whitney test.

The conceptual contribution of the study lies in treating AI capability and forensic-accounting expertise as complementary rather than substitutive elements of tax-crime enforcement. AI can screen large transactional datasets, identify anomalous relationships, reveal networks of related entities, compare tax and third-party information, and rank taxpayers or transactions according to estimated risk. As result, these outputs remain investigative signals.

Once suspected fraud are identified in the investigative stage, its economic and financial impact must be calculated by forensic accountants by the reconstruction of transaction flow, tracing of payment flows, examination of accounting records, identification of related parties, quantification of fiscal damage, evaluation of alternative explanations, and corroboration of findings through independent evidence.

This interpretation suggests that increasing technological capability need not reduce the importance of forensic accounting. On the contrary, improved detection may increase the number or complexity of cases requiring downstream professional investigation. The data are compatible with this interpretation but do not directly demonstrate it. What they show is that AI-implementing countries exhibit greater revenue-fraud investigative activity and, most consistently, a greater concentration of revenue-fraud cases within EPPO caseloads, but the dataset cannot determine whether forensic-accounting staffing, capability, or actual professional participation increased in parallel.

The findings complement existing studies demonstrating the predictive or operational benefits of AI. Battaglini et al. (2025), for example, demonstrate that machine-learning approaches can improve tax-audit targeting. Belahouaoui and Attak (2024), Anjarwi (2026), and Zaim and Sahbani (2026) likewise identify substantial potential for advanced analytics in tax compliance and fraud detection. The present study examines a later stage of the enforcement process.

Conclusion

This study examines whether AI implementation in EU tax administrations is associated with greater volume of investigative activity in revenue-fraud cases requiring substantial forensic-financial analysis. Data used in the study combines country-level AI implementation information from ISORA with EPPO revenue-fraud investigation statistics. The analytical sample comprises 24 EPPO-

participating EU Member States and are 21 classified as AI implementers and three as non-implementers. Three complementary indicators capture the volume, prosecutorial concentration, and financial materiality of revenue-fraud investigations. Across all three, AI-implementing Member States exhibit higher means and medians than non-implementers.

The strongest and most consistent evidence concerns RFIS, the share of revenue-fraud investigations within each country's total active EPPO caseload. AI-implementing countries recorded an average RFIS of 32,75%, compared with 13,25% for non-implementers. The difference is statistically significant under the Welch test ($t = 3,619, p = 0,004$) and remains significant in the OLS specification ($b = 15,54, p = 0,007, R^2 = 0,416$). The Mann-Whitney result ($U = 52,5, p = 0,074$) provides weaker, marginal evidence at the 10% level, suggesting that the difference is directionally consistent across the distribution but not sufficiently strong to meet the conventional 5% threshold. By contrast, the per-capita investigation rate (RFIR) and the financial-materiality measure (RFFM) are significant under the Welch test and OLS regression but not under the Mann-Whitney test ($p = 0,452$ and $p = 0,145$, respectively). This divergence indicates that the RFIR and RFFM results are more sensitive to the magnitude of individual observations, whereas RFIS provides the most stable evidence of a systematic difference between AI-implementing and non-implementing countries. The findings are therefore consistent with the proposed hypothesis, but they do not constitute strong or uniform confirmation.

The study's principal contribution lies in connecting AI-enabled tax administration with the volume of forensic-investigative activity. AI can increase the capacity to identify suspicious transactions, anomalous relationships, and potentially complex fraud structures, but these signals must subsequently be reconstructed, quantified, corroborated, and developed into evidence capable of supporting criminal investigation and prosecution. In this sense, AI and forensic accounting are complementary capabilities. As result, AI contributes primarily to identification and prioritization, and forensic accounting contributes to analysis, interpretation, financial quantification, and evidentiary substantiation.

The forensic-accounting interpretation nevertheless remains indirect. EPPO revenue-fraud cases are used for estimating forensic-accounting investigative activity rather than as a direct measure of forensic-accountant participation.

Future research should therefore cover several directions. For instance, an AI-maturity indicator could distinguish between risk assessment, audit targeting, fraud detection, network analysis, taxpayer services, and other applications. Also, longitudinal data would permit within-country analysis of whether changes in AI use precede changes in the volume or composition of revenue-fraud investigations. Most importantly, comparable information on the direct involvement of forensic accountants or specialized financial investigators would allow the activity-based proxy used in this study to be validated. Such data could reveal whether AI increases

the number of cases requiring forensic-accounting intervention, changes their complexity, reduces investigation time, affects the amount of quantified loss, or influences the probability of successful prosecution.

References

- Adan, H., Atsebi, J.-M., Gueorguiev, N., Honda, J., & Nose, M. (2023). *Quantifying the revenue yields from tax administration reforms* (IMF Working Paper No. 23/231). International Monetary Fund.
- Adelakun, B. O., Onwubuariri, E. R., Adeniran, G. A., & Ntiakoh, A. (2024). Enhancing fraud detection in accounting through AI: Techniques and case studies. *Finance & Accounting Research Journal*, 6(6), 978–999. <https://doi.org/10.51594/farj.v6i6.1232>
- Anjarwi, A. W. (2026). The digital transformation of tax audits: How AI, big data, blockchain, and advanced analytics are reshaping tax evasion detection. *Journal of Business Analytics*. Advance online publication. <https://doi.org/10.1080/2573234X.2026.2644363>
- Battaglini, M., Guiso, L., Lacava, C., Miller, D. L., & Patacchini, E. (2025). Refining public policies with machine learning: The case of tax auditing. *Journal of Econometrics*, 249, Article 105847.
- Belahouaoui, R., & Attak, E. H. (2024). Digital taxation, artificial intelligence and Tax Administration 3.0: Improving tax compliance behavior—A systematic literature review using textometry (2016–2023). *Accounting Research Journal*, 37(2), 172–191. <https://doi.org/10.1108/ARJ-12-2023-0372>
- Buil-Gil, D., Pina-Sánchez, J., & Aebi, M. F. (2026). The effect of counting rules on cross-national comparisons of homicide. *Social Indicators Research*, 181(1), Article 44. <https://doi.org/10.1007/s11205-025-03802-8>
- De Simone, L., Stomberg, B., & Williams, B. (2023). Does tax enforcement disparately affect domestic versus multinational corporations around the world? *Contemporary Accounting Research*, 40(4), 2413–2445. <https://doi.org/10.1111/1911-3846.12900>
- European Public Prosecutor's Office. (2026). *EPPO annual report 2025*. Publications Office of the European Union. <https://doi.org/10.2927/2362576>
- Eurostat. (2026). *Main national accounts tax aggregates: gov_10a_taxag* [Data set]. European Commission.
- Judijanto, L., & Nurhayati, S. (2026). Artificial intelligence in tax administration: A review of implementation strategies, challenges, and performance outcomes. In *Proceedings of the 2026 9th International Conference on Information and Computer Technologies (ICICT 2026)*. Association for Computing Machinery. <https://doi.org/10.1145/3803291.3803330>

- Kaufmann, D., & Kraay, A. (2024). *The Worldwide Governance Indicators: Methodology and 2024 update* (World Bank Policy Research Working Paper No. 10952). World Bank.
- Lyeonov, S., Draskovic, V., Kubaščíková, Z., & Fenyves, V. (2024). Artificial intelligence and machine learning in combating illegal financial operations: Bibliometric analysis. *Human Technology*, 20(2), 325–360. <https://doi.org/10.14254/1795-6889.2024.20-2.5>
- Meneses Cerón, L. Á., Fernández Rengifo, E., Pino Terán, V. E., Albán Lara, M. A., Vega Sánchez, R. C., & Fuentes Sánchez, J. J. (2026). Artificial intelligence and machine learning in the transformation of forensic accounting: A systematic literature review. *Journal of Risk and Financial Management*, 19(7), Article 519. <https://doi.org/10.3390/jrfm19070519>
- Milosavljević, M., Ignjatović, M., Spasenić, Ž., Milanović, N., & Đoković, A. (2024). Is there a link between tax administration performance and tax evasion? *Economies*, 12(8), Article 193. <https://doi.org/10.3390/economies12080193>
- OECD. (2025a). *Tax administration 2025: Comparative information on OECD and other advanced and emerging economies*. OECD Publishing. <https://doi.org/10.1787/cc015ce8-en>
- OECD. (2025b). *Governing artificial intelligence: AI in tax administration*. OECD Publishing.
- OECD. (2025c). *Inventory of Tax Technology Initiatives (ITTI): Data Explorer*. OECD.
- Qatawneh, A. (2024). The role of artificial intelligence in auditing and fraud detection in accounting information systems: Moderating role of natural language processing. *International Journal of Organizational Analysis*, 33(6), 1391–1409.
- Thakkar, H., Datta, S., Bhadra, P., Barot, H., & Jadav, J. (2025). Artificial intelligence and machine learning in fraud detection: A comprehensive bibliometric mapping of research trends and directions. *Annals of Library and Information Studies*, 72, 138–150. <https://doi.org/10.56042/alis.v72i2.14752>
- World Bank. (2025a). *World Development Indicators*. World Bank.
- World Bank. (2025b). *Worldwide Governance Indicators*. World Bank.
- Zaim, H., & Sahbani, S. (2026). Artificial intelligence in tax compliance and evasion mitigation: Trends, mechanisms, and institutional implications. *Journal of Risk and Financial Management*, 19(7), Article 513. <https://doi.org/10.3390/jrfm19070513>

Sofija Sekulić
Faculty of Economics, University of
Montenegro, Podgorica
ORCID: 0009-0003-4143-8455

Ana Lalević Filipović
Faculty of Economics, University of
Montenegro, Podgorica
ORCID: 0009-0007-3595-6570

INSTITUTIONAL ENVIRONMENT AS A PREREQUISITE FOR THE DEVELOPMENT OF FORENSIC ACCOUNTING: A CLUSTER AND COMPARATIVE ANALYSIS OF EUROPEAN UNION AND WESTERN BALKAN COUNTRIES

UDK 657.6:330.34

DOI <https://doi.org/10.54318/czf.2026.ss.009>

Abstract

Forensic accounting has long been recognized as one of the key mechanisms for the prevention, detection, and investigation of financial fraud, corruption, money laundering, and other forms of economic crime. Its importance has grown steadily in both professional practice and academic research. However, the effective application of forensic accounting depends not only on the expertise of forensic accountants but also on the quality of the institutional environment in which investigative, regulatory, and judicial authorities operate. Although the literature increasingly emphasizes the importance of institutional frameworks for conducting effective financial investigations, studies that systematically examine institutional characteristics as a prerequisite for the development and effective application of forensic accounting remain relatively limited.

Against this background, the aim of this paper is to identify the institutional profiles of European Union and Western Balkan countries and to analyse their similarities and differences in terms of the quality of the institutional environment relevant to the development of forensic accounting. The empirical analysis is based on the Worldwide Governance Indicators (WGI) of the World Bank for 2024, using six governance dimensions: Voice and Accountability, Political Stability and Absence of Violence/Terrorism, Government Effectiveness, Regulatory Quality, Rule of Law, and Control of Corruption. The sample comprises 32 countries – the 27 European Union member states and five Western Balkan countries (Albania, Bosnia and Herzegovina, Montenegro, North Macedonia, and Serbia) for which complete data are available. Hierarchical cluster analysis (Ward's method) is employed to identify homogeneous groups of countries with similar institutional characteristics, complemented by k-means clustering and non-parametric comparative analysis. The results reveal a three-tier institutional architecture of the European economic space, in which all Western Balkan countries, together with Bulgaria, Romania, Hungary, and Greece, form the institutionally weakest cluster. The EU–Western Balkans gap is largest precisely in the dimensions most directly relevant to forensic accounting:

Control of Corruption (1.37 points) and Rule of Law (1.21 points). The findings contribute to a better understanding of the institutional prerequisites for the development of forensic accounting, identify the key institutional challenges faced by the Western Balkan countries, and provide recommendations for strengthening institutional frameworks that support more effective financial investigations and the fight against financial fraud and corruption.

Key words: *Forensic accounting, Institutional environment, Cluster analysis, Comparative analysis, Worldwide Governance Indicators, European Union, Western Balkans.*

JEL classification: *M41, M42, K42, O52*

Introduction

Financial fraud, corruption, and money laundering impose substantial economic and social costs on modern economies. According to the Association of Certified Fraud Examiners, organizations lose an estimated five percent of their annual revenues to occupational fraud, with median losses that are systematically higher in environments with weaker oversight and control mechanisms (ACFE, 2024). In response to the growing scale and sophistication of economic crime, forensic accounting has developed into a distinct professional and academic discipline that integrates accounting, auditing, and investigative skills for the purpose of preventing, detecting, and investigating financial irregularities, and for providing expert evidence in judicial proceedings (Crumley, Heitger & Smith, 2017; Honigsberg, 2020).

A substantial body of research has examined the competencies of forensic accountants, the techniques of fraud detection, and the organizational determinants of forensic accounting practice (DiGabriele, 2008; Rezaee, 2005; Bhasin, 2007). Considerably less attention has been devoted to the macro-institutional preconditions under which forensic accounting can function effectively. Yet the discipline is, by its very nature, institutionally embedded: the demand for forensic accounting services depends on the willingness and capacity of regulators, prosecutors, and courts to pursue financial crime; the evidentiary value of forensic findings depends on the independence and quality of the judiciary; and the deterrent effect of financial investigations depends on the credibility of enforcement. Where institutions are captured, politicized, or ineffective, even technically excellent forensic work may fail to translate into accountability. In the terminology of institutional economics, forensic accounting is a mechanism of third-party enforcement whose effectiveness is conditional on the broader institutional matrix within which it operates (North, 1990; Acemoglu & Robinson, 2012).

This institutional perspective is particularly salient for the Western Balkan countries. All of them are either candidates for, or in an advanced process of, accession to the European Union, and the accession framework itself – notably the 'fundamentals

first' approach centred on Chapters 23 and 24 of the *acquis* – places the rule of law, judicial reform, and the fight against corruption and organized crime at the core of the enlargement process (European Commission, 2024). Persistent deficits in judicial independence, prosecutorial capacity, and anti-corruption enforcement have repeatedly been identified as the main bottlenecks of the region's European integration. These same deficits define the environment in which forensic accountants in the region must operate.

Against this background, the aim of this paper is twofold. First, it seeks to identify the institutional profiles of European Union and Western Balkan countries using the six dimensions of the World Bank's Worldwide Governance Indicators (WGI) for 2024, applying cluster analysis to detect homogeneous groups of countries with similar institutional characteristics. Second, it uses comparative analysis to establish the position of the Western Balkan countries relative to the EU member states and to interpret the resulting institutional configurations from the perspective of forensic accounting development. The paper addresses three research questions:

RQ1: What institutional profiles can be identified among EU and Western Balkan countries on the basis of the WGI dimensions?

RQ2: How large and how systematic is the institutional gap between the Western Balkans and the EU, and in which governance dimensions is it most pronounced?

RQ3: What do the identified institutional configurations imply for the development and effective application of forensic accounting in the region?

The remainder of the paper is organized as follows. Section 1 reviews the literature on forensic accounting and its institutional determinants. Section 2 describes the data and methodology. Section 3 presents the results of the descriptive, comparative, and cluster analyses. Section 4 discusses the findings and their implications for forensic accounting, while final Section concludes with recommendations, limitations, and directions for future research.

1. Literature Review

1.1. Forensic accounting: concept and development

Forensic accounting is commonly defined as the application of accounting, auditing, and investigative skills to matters that are anticipated to be resolved in a legal context (Crumbley et al., 2017). It encompasses fraud examination, litigation support, business valuation in disputes, and the tracing of illicit financial flows. Honigsberg (2020) emphasizes that forensic accounting occupies the intersection of accounting and law, and that its societal value derives from its capacity to produce credible evidence that can withstand adversarial scrutiny. Rezaee (2005) situates forensic accounting within the broader anti-fraud architecture, arguing that financial statement fraud arises from the interaction of incentives, opportunities, and weak governance, and that effective deterrence requires both firm-level controls and

credible external enforcement. DiGabriele (2008) documents the specific skill set: analytical proficiency, investigative intuition, and courtroom communication, that distinguishes forensic accountants from traditional auditors, while Bhasin (2007) and Huber (2012) trace the professionalization of the discipline and the emergence of dedicated certifications and professional bodies.

In transition and post-transition economies, the development of forensic accounting has been slower and more uneven. The literature on South-East Europe points to a limited number of certified practitioners, the absence or recency of national certification schemes, weak integration of forensic accounting into university curricula, and, crucially, limited institutional demand for forensic expertise from prosecutors, courts, and regulators. In such settings, forensic accounting exists more as a potential than as an operating mechanism of accountability, which directs analytical attention precisely to the institutional environment examined in this paper.

1.2. The institutional environment as a determinant of forensic accounting

The theoretical link between institutional quality and forensic accounting can be articulated through three channels. The first is the demand channel: forensic accounting services are demanded when there is a realistic prospect that financial crime will be investigated and prosecuted. Government effectiveness and control of corruption determine whether investigative and regulatory bodies possess the capacity, resources, and integrity to initiate financial investigations. The second is the enforcement channel: the rule of law – the quality of contract enforcement, property rights, the police, and the courts, determines whether forensic findings can be transformed into convictions, asset recoveries, and sanctions. La Porta et al. (1999) show that the quality of government systematically shapes the performance of legal and regulatory institutions across countries. The third is the transparency channel: voice and accountability, including media freedom and civil society scrutiny, generate the informational environment in which irregularities are surfaced, whistle-blowers are protected, and public pressure for investigations is sustained.

Conversely, weak institutions undermine each link of this chain. Where corruption penetrates the judiciary and the prosecution, forensic evidence may be ignored or suppressed; where regulatory quality is low, the legal predicates for financial investigations (accounting, auditing, anti-money-laundering legislation) are incomplete or inconsistently applied; where political stability is fragile, enforcement priorities shift with political cycles. It follows that the institutional environment may constitute an important structural prerequisite for the development and effective application of forensic accounting. Accordingly, this study examines the institutional conditions that may facilitate or constrain such development, rather than directly measuring the level of forensic accounting development itself.

1.3. Measuring institutional quality: the Worldwide Governance Indicators

The Worldwide Governance Indicators, developed by Kaufmann, Kraay, and Mastruzzi, are among the most widely used composite measures of institutional quality, covering more than 200 economies since 1996 (Kaufmann, Kraay & Mastruzzi, 2011; Kaufmann & Kraay, 2024). The WGI aggregate perception data from 35 cross-country sources into six dimensions: Voice and Accountability (VA), Political Stability and Absence of Violence/Terrorism (PS), Government Effectiveness (GE), Regulatory Quality (RQ), Rule of Law (RL), and Control of Corruption (CC). The 2025 revision of the WGI introduced methodological refinements to source screening, indicator mapping, and the aggregation model, with historical estimates recalculated back to 1996 to ensure consistency over time (World Bank, 2025). Although the WGI have been debated with respect to perception bias and cross-source correlation, they remain the standard instrument for cross-country institutional comparison and are particularly suitable for the classification exercise undertaken here, since all six dimensions are conceptually linked to the functioning of forensic accounting.

2. Data and Methodology

2.1. Sample and data

The sample comprises 32 European countries: the 27 member states of the European Union and five Western Balkan countries: Albania, Bosnia and Herzegovina, Montenegro, North Macedonia, and Serbia. Data refer to the year 2024 and are taken from the Worldwide Governance Indicators of the World Bank (2025 revision), which is the first WGI release to report estimates for 2024 (World Bank, 2025). The governance estimates, expressed in units of a standard normal distribution (approximately -2.5 to $+2.5$), are used rather than percentile ranks, since the estimates are better suited to distance-based multivariate methods.

Table 1 summarizes the six WGI dimensions and their relevance for forensic accounting.

Table 1: WGI dimensions and their relevance for forensic accounting

Dimension	What it measures	Relevance for forensic accounting
Voice and Accountability (VA)	Participation in selecting government; freedom of expression, association, and media	Media and civil-society scrutiny that surfaces irregularities; whistle-blower environment
Political Stability (PS)	Likelihood of political instability and politically motivated violence	Continuity of enforcement priorities; insulation of investigations from political shocks
Government Effectiveness (GE)	Quality of public services, civil service, and policy implementation	Capacity and resources of investigative, tax, and regulatory bodies
Regulatory Quality (RQ)	Ability of government to formulate and implement sound regulations	Quality of accounting, auditing, AML, and market-integrity legislation
Rule of Law (RL)	Confidence in and abidance by rules; contract enforcement, police, and courts	Judicial follow-through on forensic evidence; convictions and asset recovery
Control of Corruption (CC)	Extent to which public power is exercised for private gain; state capture	Integrity of the enforcement chain; demand for, and independence of, financial investigations

Source: Authors' compilation based on World Bank (2025)

2.2. Methods

The empirical strategy combines three complementary techniques. First, descriptive and comparative statistics are computed for the EU-27 and the Western Balkan group, and the significance of group differences is assessed with the non-parametric Mann–Whitney U test, which is appropriate given the small size of the Western Balkan group and the absence of distributional assumptions. Second, the interdependence of the six dimensions is examined through Pearson correlations and principal component analysis. Third, cluster analysis is applied to identify homogeneous institutional profiles. Prior to clustering, all variables are standardized

(z-scores) to remove scale effects. Hierarchical agglomerative clustering with Ward's minimum-variance criterion and squared Euclidean distances is used as the primary method, as it is the standard choice for producing compact, interpretable groups in socio-economic classification (Ward, 1963; Everitt et al., 2011). The number of clusters is selected on the basis of the dendrogram structure, average silhouette widths for $k = 2-6$ (Milligan & Cooper, 1985), and the substantive interpretability of the solutions. The robustness of the partition is verified by re-estimating the classification with the k-means algorithm (50–100 random initializations) and computing the Adjusted Rand Index between the two solutions, and by inspecting the adjacent four-cluster solution.

3. Results

3.1. Descriptive and comparative analysis

Table 2 reports the 2024 WGI estimates for all 32 countries, together with a simple composite index computed as the unweighted mean of the six dimensions. Table 3 summarizes the group comparison.

Table 2: WGI 2024 estimates for EU and Western Balkan countries

Country	Group	VA	PS	GE	RQ	RL	CC	Composite
Denmark	EU	1.89	0.77	1.91	1.93	2.00	2.29	1.80
Finland	EU	1.85	0.83	1.82	1.96	2.02	2.19	1.78
Luxembourg	EU	1.67	1.07	2.04	1.84	1.76	2.02	1.73
Sweden	EU	1.73	0.60	1.77	1.67	1.73	1.97	1.58
Netherlands	EU	1.65	0.43	1.75	1.72	1.71	1.91	1.53
Ireland	EU	1.64	0.71	1.61	1.79	1.59	1.57	1.49
Germany	EU	1.60	0.12	1.55	1.50	1.63	1.77	1.36
Estonia	EU	1.41	0.70	1.30	1.40	1.49	1.67	1.33
Austria	EU	1.41	0.49	1.51	1.35	1.68	1.29	1.29
Lithuania	EU	1.22	0.94	1.20	1.29	1.37	0.99	1.17
Belgium	EU	1.50	0.12	1.19	1.24	1.28	1.47	1.13
Czechia	EU	1.25	0.97	1.18	1.23	1.19	0.93	1.12
Latvia	EU	1.13	0.62	0.74	1.10	1.09	0.83	0.92
France	EU	1.15	-0.24	1.20	1.15	0.98	1.22	0.91
Portugal	EU	1.27	0.54	0.95	0.74	1.07	0.82	0.90
Slovenia	EU	1.06	0.72	1.09	0.80	0.98	0.67	0.89
Spain	EU	1.15	0.00	1.11	0.62	0.93	0.71	0.75
Cyprus	EU	0.89	0.39	0.87	0.74	0.69	0.53	0.68
Italy	EU	1.07	0.31	0.79	0.64	0.62	0.55	0.66
Poland	EU	0.75	0.50	0.67	0.69	0.58	0.79	0.66
Malta	EU	0.92	0.80	0.62	0.49	0.73	0.27	0.64
Slovakia	EU	0.93	0.63	0.66	0.40	0.56	0.11	0.55
Croatia	EU	0.65	0.58	0.71	0.29	0.38	0.15	0.46

Country	Group	VA	PS	GE	RQ	RL	CC	Composite
Greece	EU	0.78	0.14	0.19	0.48	0.30	0.30	0.36
Romania	EU	0.39	0.16	0.42	0.38	0.37	0.03	0.29
Hungary	EU	-0.03	0.44	0.50	0.34	0.19	0.04	0.25
Montenegro	WB	0.39	0.33	0.37	0.13	0.18	-0.06	0.22
Bulgaria	EU	0.21	0.04	0.04	0.29	-0.03	-0.24	0.05
Albania	WB	-0.01	0.02	0.31	0.14	-0.12	-0.42	-0.01
North Macedonia	WB	0.07	0.16	-0.03	0.27	-0.22	-0.33	-0.01
Serbia	WB	-0.47	-0.17	0.10	0.08	-0.20	-0.56	-0.20
Bosnia and Herzegovina	WB	-0.40	-0.50	-0.64	-0.35	-0.32	-0.51	-0.45

Source: Authors' compilation

Table 3: Group comparison: EU-27 vs Western Balkans

Dimension	EU mean	EU s.d.	WB mean	WB s.d.	Gap (EU-WB)	Mann-Whitney p
Voice and Accountability	1.153	0.489	-0.084	0.355	1.237	0.001
Political Stability	0.496	0.328	-0.032	0.320	0.528	0.010
Government Effectiveness	1.089	0.536	0.022	0.403	1.067	0.001
Regulatory Quality	1.040	0.557	0.054	0.236	0.986	0.001
Rule of Law	1.070	0.585	-0.136	0.190	1.206	0.001
Control of Corruption	0.994	0.741	-0.376	0.197	1.370	<0.001
Composite index	0.974	0.432	-0.092	0.267	1.066	<0.001

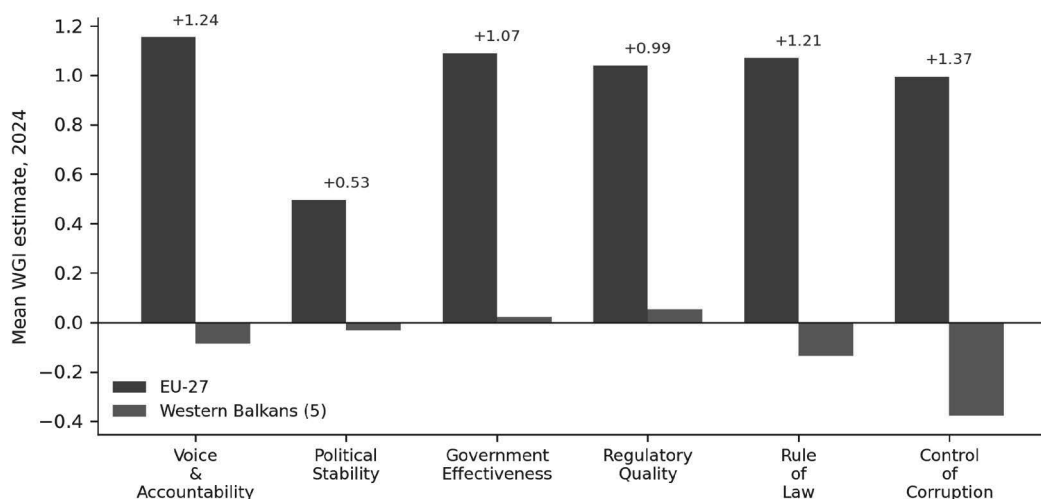
Note: two-sided Mann-Whitney U test of the difference between the EU-27 and Western Balkan groups.

Source: Authors' calculations

Three findings stand out. First, the institutional gap between the EU and the Western Balkans is statistically significant for every dimension despite the small size of the Western Balkan group, indicating that the difference is systematic rather than driven by individual countries. Second, the gap is not uniform: it is smallest for Political Stability (0.53 points) and largest precisely in the two dimensions most directly connected to the enforcement of financial accountability – Control of Corruption (1.37) and Rule of Law (1.21). All five Western Balkan countries record negative

Control of Corruption estimates, meaning that they fall below the world average on this dimension, whereas the EU mean is close to one full standard deviation above it. Third, the Western Balkan group is internally compact: its standard deviations are smaller than those of the EU on five of six dimensions, suggesting a shared regional institutional configuration rather than five idiosyncratic national situations. Figure 1 visualizes the group comparison, where labels above the bars show the EU–WB gap in points.

Figure 1: Mean WGI estimates by dimension - EU-27 vs Western Balkans



Source: Authors' calculations

At the level of individual countries (Table 2), Montenegro is the strongest Western Balkan performer, with a composite score (0.223) that places it above two EU member states - Bulgaria (0.052) and, marginally, close to Hungary (0.247), while Bosnia and Herzegovina is the weakest country in the entire sample (−0.453), recording the lowest values on four of six dimensions. Serbia displays the sharpest asymmetry in the region: its Voice and Accountability estimate (−0.47) is the lowest of all 32 countries, while its Government Effectiveness (0.10) is closer to the regional average, a configuration consistent with capable administration operating under weak democratic accountability.

3.2. Interdependence of the governance dimensions

Table 4 reports the Pearson correlation matrix. Five of the six dimensions: VA, GE, RQ, RL, and CC – are very highly intercorrelated (r between 0.92 and 0.97), while Political Stability correlates with the remaining dimensions at a distinctly lower level (r between 0.52 and 0.62). Principal component analysis confirms this structure: the first component captures 86.9% of the total variance and loads strongly on all dimensions, while the second component (10.0%) separates Political Stability from

the rest. Substantively, this indicates that institutional quality in the European economic space is essentially a single underlying dimension - a 'governance syndrome' in which accountability, state capacity, regulation, the rule of law, and corruption control move together, with political stability behaving partly independently (as illustrated by France, which combines strong institutions with the lowest Political Stability estimate in the EU in 2024). For forensic accounting, this implies that deficits in corruption control are rarely isolated: they co-occur with weaknesses across the entire enforcement chain.

Table 4: Pearson correlations among the WGI dimensions (N = 32)

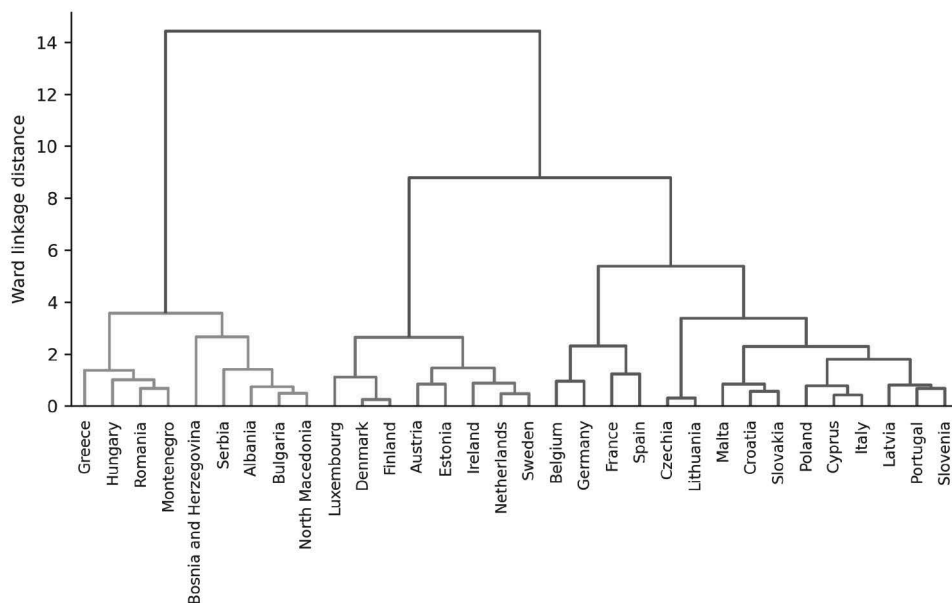
	VA	PS	GE	RQ	RL	CC
VA	1.000	0.609	0.927	0.919	0.960	0.940
PS	0.609	1.000	0.617	0.588	0.618	0.524
GE	0.927	0.617	1.000	0.947	0.961	0.946
RQ	0.919	0.588	0.947	1.000	0.964	0.971
RL	0.960	0.618	0.961	0.964	1.000	0.970
CC	0.940	0.524	0.946	0.971	0.970	1.000

Source: Authors' calculations

3.3. Cluster analysis

Figure 2 presents the dendrogram of the hierarchical clustering (Ward's method, squared Euclidean distances, standardized data). Average silhouette widths are 0.44 for $k = 2$, 0.39 for $k = 3$, and 0.41 for $k = 4$; while the two-cluster solution is marginally superior on this criterion, it merely separates a small Nordic-type elite from the rest and is analytically uninformative for positioning the Western Balkans. The three-cluster solution is retained as the primary partition because it is clearly visible in the dendrogram, substantively interpretable, and robust: the k-means re-estimation reproduces the top and bottom clusters exactly.

Figure 2: Dendrogram of the hierarchical cluster analysis



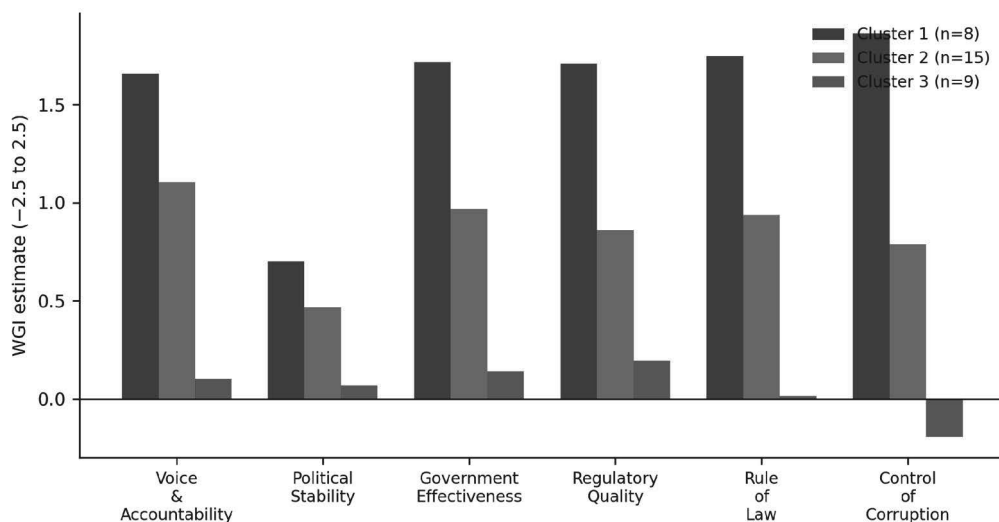
Source: Authors' calculations using Statistica software

The three clusters, labelled in descending order of institutional quality, are reported in Table 5, and their centroids are presented in Figure 3.

Table 5: Cluster membership (Ward's method, $k = 3$)

Cluster	n	Members
Cluster 1 – Institutional leaders	8	Austria, Denmark, Estonia, Finland, Ireland, Luxembourg, Netherlands, Sweden
Cluster 2 – Consolidated core	15	Belgium, Croatia, Cyprus, Czechia, France, Germany, Italy, Latvia, Lithuania, Malta, Poland, Portugal, Slovakia, Slovenia, Spain
Cluster 3 – Institutionally lagging group	9	Bulgaria, Greece, Hungary, Romania, Albania, Bosnia and Herzegovina, Montenegro, North Macedonia, Serbia

Source: Authors' calculations using Statistica software

Figure 3: Cluster profiles: mean WGI estimates by dimension

Source: Authors' calculations using Statistica software

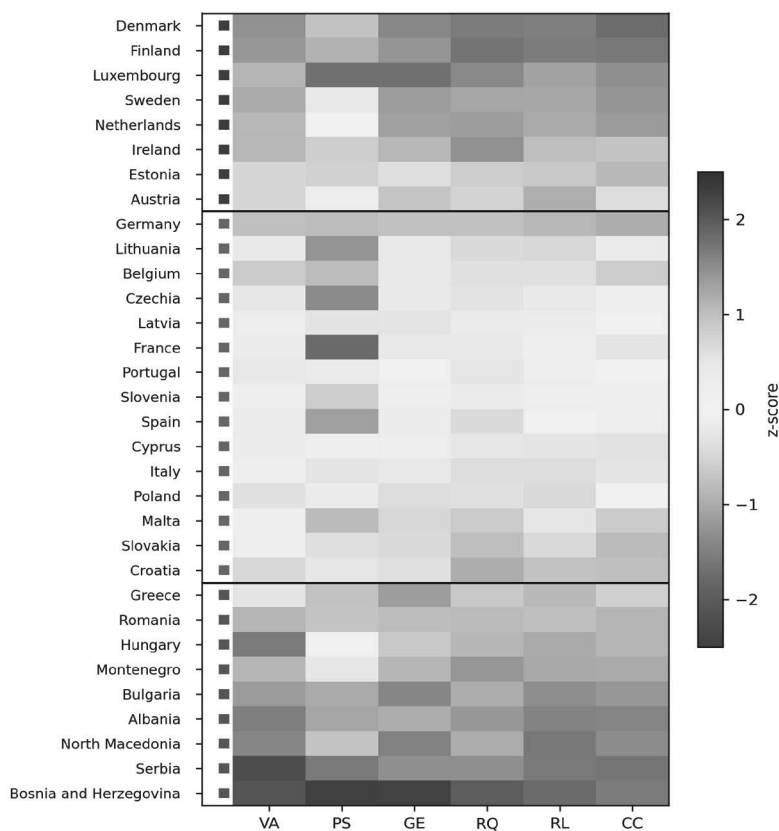
Cluster 1 comprises the Nordic member states, the Benelux core, Austria, Ireland, and, notably, Estonia, the only post-communist country in the top group. Its centroid lies between 1.6 and 1.9 points on every dimension except Political Stability, roughly one standard deviation above the sample mean throughout. Cluster 2 gathers the large continental economies, the Mediterranean member states, and the more successful Central and East European members; its profile is solidly positive but 0.6–1.1 points below the leaders, with Control of Corruption as its relatively weakest dimension. Cluster 3 unites all five Western Balkan countries with four EU member states: Bulgaria, Romania, Hungary, and Greece. Its centroid hovers around zero on five dimensions and is negative for Control of Corruption (–0.19); relative to the sample it lies more than one standard deviation below the mean on every dimension except Political Stability.

Two features of this solution deserve emphasis. First, the boundary of the weakest cluster does not coincide with the EU border: EU membership per se does not guarantee escape from the lagging institutional profile, as the presence of four member states in Cluster 3 demonstrates. Conversely, the fact that Montenegro's profile is statistically indistinguishable from that of Bulgaria or Hungary shows that parts of the Western Balkans have already converged with the weaker segment of the EU. Second, within Cluster 3 the Western Balkan countries nevertheless occupy the lower stratum: on the composite index, Bulgaria (0.05) is the only EU member below every Western Balkan country except Montenegro, while Serbia and Bosnia and Herzegovina anchor the bottom of the entire distribution.

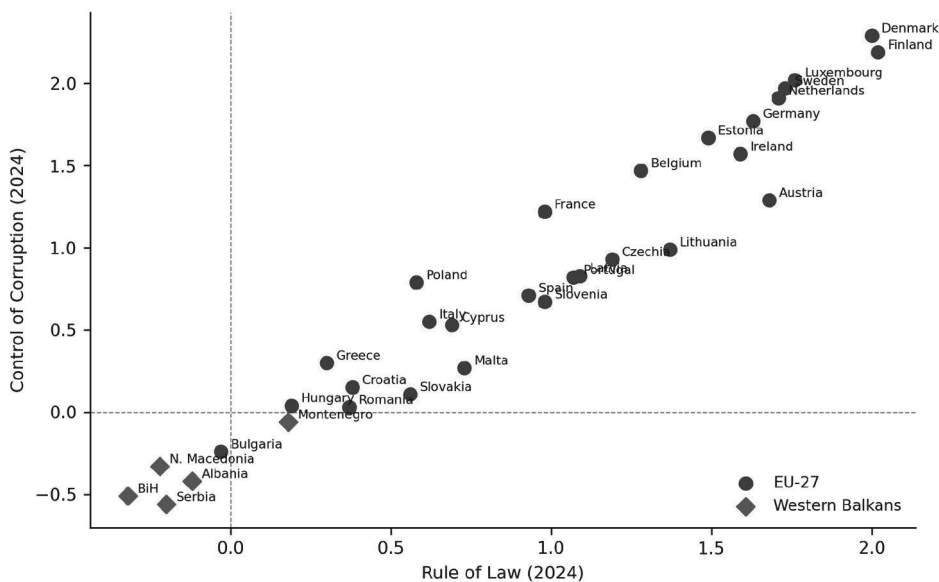
Taken together, these findings highlight both the cross-border nature of institutional weakness and the internal stratification within the lagging cluster, patterns that are

further illustrated in Figures 4 and 5. Figure 4 summarizes the full classification as a heatmap of standardized scores, where countries are sorted by cluster and composite score, horizontal lines separate clusters and the colour of the left-hand marker denotes cluster membership. Figure 5 positions the countries in the space of the two dimensions most relevant for forensic accounting - Rule of Law and Control of Corruption.

Figure 4: Standardized WGI profiles (z-scores) by cluster



Source: Authors' calculations using Statistica software

Figure 5: Rule of Law vs Control of Corruption

Source: Authors' calculations using Statistica software

4. Discussion: Implications for forensic accounting

Interpreted through the three channels developed in Section 2, the results provide insight into the institutional conditions that may facilitate or constrain the development and effective application of forensic accounting in the Western Balkans.

First, on the demand channel, the negative Control of Corruption estimates across the entire region indicate an environment in which the principal potential 'clients' of forensic accounting – prosecutors' offices, anti-corruption agencies, tax administrations, and financial-intelligence units – operate under conditions of elevated capture risk. Where enforcement institutions are themselves exposed to corruption, the demand for independent financial expertise is structurally suppressed: high-profile financial investigations are less likely to be initiated, and when initiated, less likely to be pursued to conclusion. The systemic character of this deficit is underscored by the correlation structure: in this sample, weak corruption control is almost never accompanied by strong rule of law or high government effectiveness.

Second, on the enforcement channel, the Rule of Law gap of 1.21 points implies that even competently produced forensic evidence faces an uncertain judicial afterlife. Forensic accounting generates value only insofar as its outputs are convertible into judicial outcomes – indictments, convictions, and asset recovery. Negative or near-

zero Rule of Law estimates in four of the five Western Balkan countries (Montenegro being the partial exception at 0.18) signal courts with limited independence and capacity for complex financial adjudication. This suggests that investment in forensic accounting skills alone, without parallel strengthening of specialized economic-crime chambers and prosecutorial financial expertise, is likely to yield low returns.

Third, on the transparency channel, the region's most heterogeneous results appear under Voice and Accountability, where Serbia (−0.47) records the weakest value in the entire 32-country sample while Montenegro (0.39) exceeds two EU member states. Since investigative journalism and civil-society scrutiny are, in weak-enforcement environments, frequently the initial detectors of financial irregularities, deterioration in this dimension removes one of the few functioning substitutes for official oversight and further narrows the channels through which forensic expertise enters public life.

Fourth, the cluster results reframe the policy narrative of 'European standards'. The relevant benchmark for the Western Balkans is not an undifferentiated EU average: the EU itself is institutionally three-tiered, and the accession frontier runs through Cluster 3, which the region shares with four member states. Two consequences follow. On the one hand, convergence with the weakest EU tier is demonstrably attainable - Montenegro has effectively reached it, and Estonia's presence in Cluster 1 shows that a post-communist starting point does not preclude reaching the institutional frontier within a generation. On the other hand, the experience of Bulgaria, Romania, and Hungary shows that entering the EU with a Cluster 3 profile does not automatically trigger further convergence; institutional quality may stagnate or regress after accession. For forensic accounting, this means that the accession process itself, with its Chapter 23/24 conditionality, monitoring, and technical assistance, is a window of opportunity during which the institutional infrastructure of financial investigation (specialized prosecution, asset-recovery offices, certified forensic expertise, protected whistle-blowing) should be built and made irreversible, rather than deferred to the post-accession period.

Conclusion

This paper set out to examine the institutional environment as a prerequisite for the development of forensic accounting, using the 2024 Worldwide Governance Indicators for the 27 EU member states and five Western Balkan countries. Three main conclusions emerge. First, cluster analysis identifies a robust three-tier institutional architecture of the European economic space - institutional leaders, a consolidated core, and an institutionally lagging group, in which all Western Balkan countries are located in the weakest cluster together with Bulgaria, Romania, Hungary, and Greece (RQ1). Second, the EU–Western Balkans institutional gap is statistically significant on all six dimensions but is deepest exactly where it matters most for forensic accounting: Control of Corruption (1.37 points) and Rule of Law

(1.21 points), while the region is internally homogeneous, indicating a shared structural configuration rather than isolated national weaknesses (RQ2). Third, the identified configuration suggests that institutional conditions may represent an important constraint on the development and effective application of forensic accounting in the region. In particular, weaknesses in corruption control, rule of law, and accountability may limit institutional demand for forensic expertise and its effective integration into enforcement processes (RQ3).

The policy recommendations follow directly. Priority should be given to: (i) building specialized institutional demand for forensic expertise: economic-crime prosecution units, financial-investigation capacity within police and tax administrations, and operationally independent anti-corruption bodies; (ii) strengthening the judicial link of the chain through specialized adjudication of complex financial cases and systematic use of court financial experts; (iii) institutionalizing the profession itself through national certification of forensic accountants, its integration into university curricula, and continuous professional education aligned with international standards; and (iv) protecting the transparency channel, media freedom and whistleblower protection, as a complementary detection mechanism. Sequencing matters: the analysis suggests these reforms deliver the greatest and most durable returns when locked in during the accession process.

The study has limitations that define the agenda for future research. The analysis is cross-sectional and based on perception-based composite indicators; it captures institutional configurations, not causal effects. Importantly, the study does not directly measure the level of forensic accounting development across countries. The WGI indicators are used to characterize the broader institutional environment that may facilitate or constrain its development and effective application. Future work should extend the analysis longitudinally to trace institutional convergence or divergence over the accession period, link the WGI profiles to direct measures of forensic accounting development (number of certified practitioners, forensic engagements, prosecuted financial-crime cases, asset-recovery outcomes), and test formally whether institutional quality Granger-causes the growth of the forensic accounting profession in transition economies.

References

- Acemoglu, D., & Robinson, J. A. (2012). *Why Nations Fail: The Origins of Power, Prosperity, and Poverty*. New York: Crown Publishers.
- Association of Certified Fraud Examiners (ACFE). (2024). *Occupational Fraud 2024: A Report to the Nations*. Austin, TX: ACFE.
- Bhasin, M. L. (2007). Forensic accounting: A new paradigm for niche consulting. *The Chartered Accountant*, 55(7), 1000–1010.
- Crumpley, D. L., Heitger, L. E., & Smith, G. S. (2017). *Forensic and Investigative Accounting* (8th ed.). Chicago: Wolters Kluwer/CCH.

DiGabriele, J. A. (2008). An empirical investigation of the relevant skills of forensic accountants. *Journal of Education for Business*, 83(6), 331–338.

European Commission. (2024). *2024 Communication on EU Enlargement Policy* (Enlargement Package). Brussels: European Commission.

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster Analysis* (5th ed.). Chichester: Wiley.

Honigsberg, C. (2020). Forensic accounting. *Annual Review of Law and Social Science*, 16, 147–164.

Huber, W. D. (2012). Is forensic accounting in the United States becoming a profession? *Journal of Forensic & Investigative Accounting*, 4(1), 255–284.

Kaufmann, D., & Kraay, A. (2024). *The Worldwide Governance Indicators: Methodology and 2024 Update*. Policy Research Working Paper. Washington, DC: World Bank Group.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2011). The Worldwide Governance Indicators: Methodology and analytical issues. *Hague Journal on the Rule of Law*, 3(2), 220–246.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (1999). The quality of government. *Journal of Law, Economics, and Organization*, 15(1), 222–279.

Milligan, G. W., & Cooper, M. C. (1985). An examination of procedures for determining the number of clusters in a data set. *Psychometrika*, 50(2), 159–179.

North, D. C. (1990). *Institutions, Institutional Change and Economic Performance*. Cambridge: Cambridge University Press.

Rezaee, Z. (2005). Causes, consequences, and deterrence of financial statement fraud. *Critical Perspectives on Accounting*, 16(3), 277–298.

Transparency International. (2025). *Corruption Perceptions Index 2024*. Berlin: Transparency International.

Ward, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, 58(301), 236–244.

World Bank. (2025). *The Worldwide Governance Indicators: Revised Methodology for Measuring Governance Using Perception Data* (2025 Revision). Washington, DC: World Bank Group. Retrieved from www.govindicators.org.

Ivana Barbieri
University of Rijeka, Faculty of
Economics and Business
ORCID: 0000-0002-3788-0975

Josip Čičak
University of Rijeka, Faculty of
Economics and Business
ORCID: 0000-0001-6090-7278

FORENSIC ACCOUNTING IN SUSTAINABILITY REPORTING: A BIBLIOMETRIC REVIEW

UDK 657.6:502.131.1

DOI <https://doi.org/10.54318/czf.2026.ib.010>

Abstract

Traditionally, financial statements have been considered the most important tool for assessing a company's business performance. Over the years, it has been observed that financial statements may be subject to manipulation aimed at presenting a misleading picture of a company's financial position and business performance. Therefore, forensic accounting has been developed as a specialised field of the accounting profession in order to identify intentional errors and manipulations of accounting information.

However, over the last 30 years, it has become clear that financial statements are insufficient for assessing the stability and long-term sustainability of business operations, and an initiative emerged to introduce additional – non-financial information. The objective of sustainability reporting is to inform stakeholders about the environmental, social, and governance aspects of business operations (ESG). Sustainability reporting includes both quantitative and qualitative information, which may also be subject to manipulation, thereby undermining users' trust in sustainability information and ultimately leading to incorrect decisions. This suggests potential for extending forensic accounting approaches to non-financial reporting.

Given that forensic accounting is associated with financial statements, most of the literature examines forensic accounting in the context of financial fraud, and its application in the field of sustainability reporting remains underexplored. Therefore, the aim of this paper is to provide a systematic literature review and to analyse whether there are studies examining the relationship between forensic accounting and sustainability reporting. The paper summarises the existing theoretical and empirical contributions and identifies the dominant research themes and contemporary trends.

This paper applies bibliometric analysis with the aim of summarising the existing scientific knowledge on the relationship between forensic accounting and

sustainability reporting. The literature search was conducted in four databases: Web of Science (WoS), Scopus, ProQuest, and The Lens.

The results show that the intersection between forensic accounting and sustainability reporting remains an underexplored research area. Existing studies suggest that forensic accounting may play a significant role in strengthening the integrity of sustainability reporting by detecting data manipulation and reducing the risk of greenwashing.

Key words: *bibliometric analysis; forensic accounting; sustainability reporting; ESG reporting; greenwashing*

JEL Classification: *M41*

Introduction

Forensic accounting plays an important role in addressing fraud and manipulation that undermine the credibility of financial information used by stakeholders to assess a company's financial position and performance (Budimir, 2017). The expansion of corporate reporting to include sustainability information raises questions about the potential application of forensic approaches beyond financial statements.

Over the last approximately 30 years, it has become clear that financial statements are insufficient for assessing the stability and long-term sustainability of business operations, and an initiative emerged to introduce additional – non-financial information. The discussion on non-financial information, i.e., non-financial reporting, gained importance in the 1990s when the processes of globalization and internationalization of companies, together with environmental problems, took place. This was known as the concept of Corporate Social Responsibility, which had already been developed in the 1950s. Since then, various standards and concepts have been developed that have attempted to unify and simplify the process of preparing non-financial reports, and today we refer to sustainability reports. Sustainability reports aim to inform stakeholders about the environmental, social, and governance aspects of business operations (ESG). In addition to quantitative information, sustainability reports also contain qualitative information, which may also be subject to manipulation, thereby undermining users' trust in the information and ultimately leading to incorrect decisions. Therefore, it is evident that the application of forensic accounting can also be extended to the area of non-financial reporting.

Although previous studies have examined ESG fraud, greenwashing, and forensic auditing, the literature remains fragmented regarding how specific forensic accounting techniques can be applied to verify sustainability disclosures (Tarjo et al., 2024; Azizah et al. 2025; Fauziah et al., 2026). In particular, empirical evidence of their effectiveness is limited, and there is a need to synthesise existing contributions, map the connections between these research streams, and identify priorities for further research. Therefore, the aim of this paper was to determine what

is known from the existing literature about the role and potential application of forensic accounting in sustainability reporting. More specifically, the intention was to determine the extent to which the relationship between sustainability reporting and forensic accounting has been investigated, which main research topics have been addressed, which types of manipulation or fraud in the field of sustainability have been examined, and to identify research gaps and opportunities for future research.

This research is motivated by the established role of forensic accounting in detecting fraud, analysing irregularities, and strengthening the credibility of financial reporting. On the other hand, sustainability and ESG reporting are increasingly becoming part of corporate reporting and an important source of information for decision-making, while at the same time concerns are growing regarding possible manipulation of ESG data and fraud in sustainability reports (Poiriazzi et al., 2025; Tarjo et al., 2024; Sun, 2025).

This paper is divided into six sections. Following the introduction, a brief presentation of the importance and application of forensic accounting in reporting is provided. This is followed by a discussion of the importance of sustainability reports and the possibility of applying forensic accounting. The fourth section presents the research process, the fifth section presents the results of the database searches, while the sixth section summarises the conclusions of the research.

1. The importance and application of forensic accounting in reporting

The fundamental role of forensic accounting is to detect, investigate, and document fraud and other irregularities, while collecting sufficient and reliable evidence on the basis of which conclusions can be drawn, recommendations can be made, or support can be provided in judicial and other legal proceedings. Therefore, a forensic accountant does not merely verify whether accounting data have been formally recorded correctly, but also investigates the circumstances in which an irregularity occurred, the persons involved, their motives, how it was carried out, and its consequences. In this sense, forensic accounting connects accounting and auditing with investigative procedures, law, information technology, and evidence analysis (Rezaee, 2019; Ozili, 2025).

Its role is both detective and preventive. In addition to detecting fraud that has already occurred, a forensic accountant assesses fraud risk, examines weaknesses in internal controls, and proposes measures aimed at reducing the possibility of irregularities recurring (Rezaee, 2019). Empirical literature also shows a positive association between the application of forensic accounting and fraud detection, as well as the identification of weak control systems (Ozili, 2025).

Traditional forensic techniques arise from a combination of accounting, auditing, and investigative procedures. The literature identifies fraud risk assessment, evaluation of internal controls, detailed analysis of financial statements and transactions, interviews and questioning of individuals, document analysis, physical

and testimonial evidence, as well as the collection and evaluation of other types of evidence (Rezaee, 2019; Ozili, 2025). In fraud risk assessment, models such as the fraud triangle, fraud diamond, and fraud pentagon are particularly used, involving the analysis of pressure/incentives, opportunity, rationalisation, capability, and other factors that create conditions for fraud (Rezaee, 2019).

Alongside traditional methods, computerised data analysis, data mining, Big Data, machine learning, artificial intelligence, and deep learning are becoming increasingly important. Their advantage lies in the ability to process large volumes of data, identify anomalies and hidden patterns, and move from retrospective fraud detection toward predictive and almost real-time fraud detection (Monserrat, 2026).

In financial reporting, the importance of forensic accounting is particularly evident because manipulated financial statements undermine transparency, information quality, and investor confidence. Forensic accounting contributes to the reliability of reports through risk assessment, analysis of internal controls, detection of material misstatements, and investigation of suspicious transactions. Rezaee (2019) identifies the reliability of financial statements as one of the areas addressed by risk-based procedures performed by forensic accountants. Bakhit (2024) provides empirical evidence that forensic accounting plays an important role in reducing financial corruption and that its effectiveness contributes to improving the quality of financial statements. Similarly, Okeke et al. (2022) identify a statistically significant relationship between forensic accounting and the integrated financial reporting of banks and conclude that its effective application has a positive impact on reporting.

A review of the forensic accounting literature and the topics examined within the field shows that forensic accounting is a multidisciplinary area, although research is not equally distributed across all of its potential domains. DiGabriele and Huber (2015) find that fraud is by far the dominant topic, while quantitative methods represent the dominant methodological approach. The authors warn that such limited diversity of topics and methods may reduce the overall contribution of forensic accounting research, since forensic accounting encompasses a much broader field than fraud examination alone. A more recent global review reaches a similar conclusion. The main themes include fraud motivation, fraud consequences, fraud detection using forensic accounting techniques, forensic accounting theory, skills, education, and employment, while the overall volume of research remains lower than in other accounting specialisations (Ozili, 2025). It is particularly noteworthy that Ozili identifies digitalization and sustainable development among the areas for future research, indicating an emerging expansion of the boundaries of the discipline.

2. The importance of sustainability reports and the potential application of forensic accounting

Alongside traditional financial statements, sustainability and ESG reporting have become important sources of information about companies' environmental and social impacts and their governance practices. Such disclosures complement

financial reporting and should contribute to greater transparency and accountability of companies toward different groups of stakeholders (Poiriazi et al., 2025). The importance of sustainability reports arises precisely from the fact that they provide users with information that cannot be fully obtained from traditional financial statements. Information on environmental impact, social responsibility, employee relations, governance, and other ESG factors enables investors and other stakeholders to gain a broader understanding of a company's operations and its long-term risks and opportunities.

Although sustainability reporting is not an entirely new phenomenon, its importance has intensified particularly over the last decade, alongside the growing interest in ESG, sustainable finance, and regulatory requirements. The increasing importance of these reports also raises the issue of their credibility. If investors, regulators, consumers, and other stakeholders increasingly rely on sustainability information when making decisions, it is not sufficient for such information merely to be disclosed; it is necessary to ensure that it is accurate, complete, comparable, verifiable, and consistent with the company's actual activities. Poiriazi et al. (2025) point out that ESG reporting still does not have the same level of rigour, controls, and verification that characterizes financial reporting. As a result, a gap may arise between what a company actually does and what it claims in its reports.

Tarjo et al. (2024) report a strong increase in ESG research from approximately 2019 onward, while Poiriazi et al. (2025) emphasize that by 2025 ESG reporting is rapidly approaching the status of a common business practice. Tarjo et al. (2024) demonstrate that fraud can no longer be viewed solely as a problem of financial statements. In their review of 66 papers on ESG fraud, non-financial reporting fraud represents the most frequently recorded ESG fraud scheme, including the failure to disclose oil spills, the concealment of child labour and poor working conditions, and the manipulation of employee data. Their analysis also demonstrates a concrete shift from traditional fraudulent financial reporting toward fraud affecting environmental and social information.

Selective disclosure, inaccurate data, and the concealment of negative events can undermine the credibility of sustainability reporting and users' trust in the disclosed information (Poiriazi et al., 2025). These risks raise the question of whether forensic accounting approaches can be adapted to verify claims concerning emissions, working conditions, resource use, and supply chains. Such verification would involve examining whether appropriate evidence supports the reported information. The following review examines how existing research has addressed these potential applications.

3. Research methodology

To determine whether there are studies investigating forensic accounting and sustainability reporting, as well as the trends in the publication of such studies, four databases were searched: Web of Science (WoS), Scopus, ProQuest, and The Lens. The databases were searched in July and August 2026. No publication date restrictions were applied. The use of the WoS and Scopus databases was justified because they provide comprehensive data from various scientific fields. Furthermore, they are recognized within the international academic community and provide access to scientific papers published in a wide range of highly ranked and highly cited journals (Harzing and Alakangas, 2016). The Google Scholar database was not searched in this case because it has been criticized as being unsuitable for bibliometric analyses due to the inclusion of outdated and inappropriate data (Harzing, 2014). ProQuest and The Lens were used as two additional databases.

The first database searched was WoS, where the search was conducted by topic (topic search). The search began with two basic terms: “forensic accounting” AND “sustainability reporting”, which resulted in 0 records. The search was then expanded using the keywords presented in Table 1. As shown in the table, when the keywords were expanded to include forensic accountant or forensic auditing, and sustainability reporting was supplemented with non-financial reporting and ESG reporting, the results again showed 0 records. Therefore, the search was further expanded by introducing the term fraud in addition to forensic accounting and auditing, which resulted in 8 records. The search was then expanded by introducing the term data manipulation, which resulted in 29 records, while the introduction of the terms greenwashing and fraud resulted in 72 records. Sustainability assurance in combination with fraud was also used in the analysis, resulting in 4 records. Therefore, a total of 113 potentially relevant records were identified in the WOS database.

Table 1: Research in WoS and Scopus

Search string	WoS Results (Number of records)	Scopus Results (Number of records)
„forensic accounting“ AND „sustainability reporting“	0	0
„forensic accounting“ OR „forensic accountant“ OR „forensic audit“ OR „forensic auditing“ AND „sustainability reporting“ OR „sustainability disclosure“ OR „non-financial reporting“ OR „nonfinancial reporting“ OR „non-financial disclosure“ OR „nonfinancial disclosure“ OR „ESG reporting“ OR „ESG disclosure“	0	2
„forensic accounting“ OR „forensic audit“ OR „forensic auditing“ AND „fraud“ or „fraud prevention“ or „fraudulent reporting“ OR „data manipulation“ AND „sustainability OR ESG OR non-financial reporting“ OR „non-financial information“	8	11
„fraud“ OR „fraud detection“ OR „fraudulent reporting“ OR „data manipulation“ OR „information manipulation“ AND „sustainability reporting“ OR „sustainability disclosure“ OR „ESG reporting“ OR „ESG disclosure“ OR „non-financial reporting“	29	51
„greenwashing“ OR „greenwashing detection“ AND „fraud“ OR „fraud detection“ OR „manipulation“	72	100
„sustainability assurance“ OR „ESG assurance“ OR „non-financial assurance“ AND „fraud“ OR „manipulation“ OR „fraud detection“	4	7

Source: Authors' research

After the WoS database, the Scopus database was searched. The search covered titles, abstracts, and keywords (TITLE-ABS-KEY). The same search strings as those used for the WoS database were applied, as shown in Table 1. As in the case of the WoS database, entering the two key terms forensic accounting and sustainability reporting resulted in 0 records. When the search was expanded by including the terms fraud and fraud detection, 51 potentially relevant records were identified. As with the WoS database search, the largest number of records was obtained when searching the keywords “greenwashing” and “fraud”, which resulted in 100 records. In total, 171 records were identified through the Scopus database search.

After searching WoS and Scopus, an additional search was conducted in the ProQuest database. In this case, the search procedure differed slightly from those used for WoS and Scopus, and different search strings were used, as shown in Table 2. Entering the selected keywords resulted in 1,999 records. The results were then restricted to documents classified as articles, books, or conference proceedings, which reduced the number of results to 185 records. The search was then further refined by specifying the following subject areas: sustainability, fraud, forensic accounting, sustainability reporting, environmental, social & governance, and fraud prevention, which ultimately resulted in 114 records.

Table 2: Research in ProQuest

Search string	Results (Number of records)	Results (after the first refinement)	Results (after the second refinement)
„forensic accounting“ OR „forensic audit“ OR „forensic auditing“ AND „sustainability reporting“ OR „sustainability disclosure“ OR „non-financial reporting“ OR „nonfinancial reporting“ OR „ESG reporting“ OR „ESG disclosure“	1999	185	114

Source: Authors' research

The final database searched was The Lens. The search strings differed slightly from those used in the previously mentioned databases, as shown in Table 3. The only condition applied was that the documents had to be journal articles. The search of this database identified 141 records. Although The Lens generated a larger number of records for the combination of terms “forensic accounting” AND “sustainability reporting” (38 records), a review of the titles and abstracts showed that many of these papers did not substantively investigate the relationship between forensic accounting and sustainability reporting.

Table 3: Research in The Lens

Search string	Results (Number of records)
„forensic accounting“ AND „sustainability reporting“	38
„forensic accounting“ AND „non-financial reporting“	8
„forensic accounting“ AND „ESG reporting“	7
„forensic accounting“ AND „greenwashing“	6
„forensic accounting“ OR „forensic auditing“ OR „forensic audit“ AND „sustainability reporting“ OR „non-financial reporting“ OR „ESG reporting“	56
„forensic accounting“ OR „forensic auditing“ AND „data manipulation“ OR „information manipulation“ OR „reporting manipulation“ AND „ESG OR sustainability OR non-financial reporting“	17
„forensic accounting“ OR „forensic auditing“ AND „greenwashing OR greenwashing detection“	9

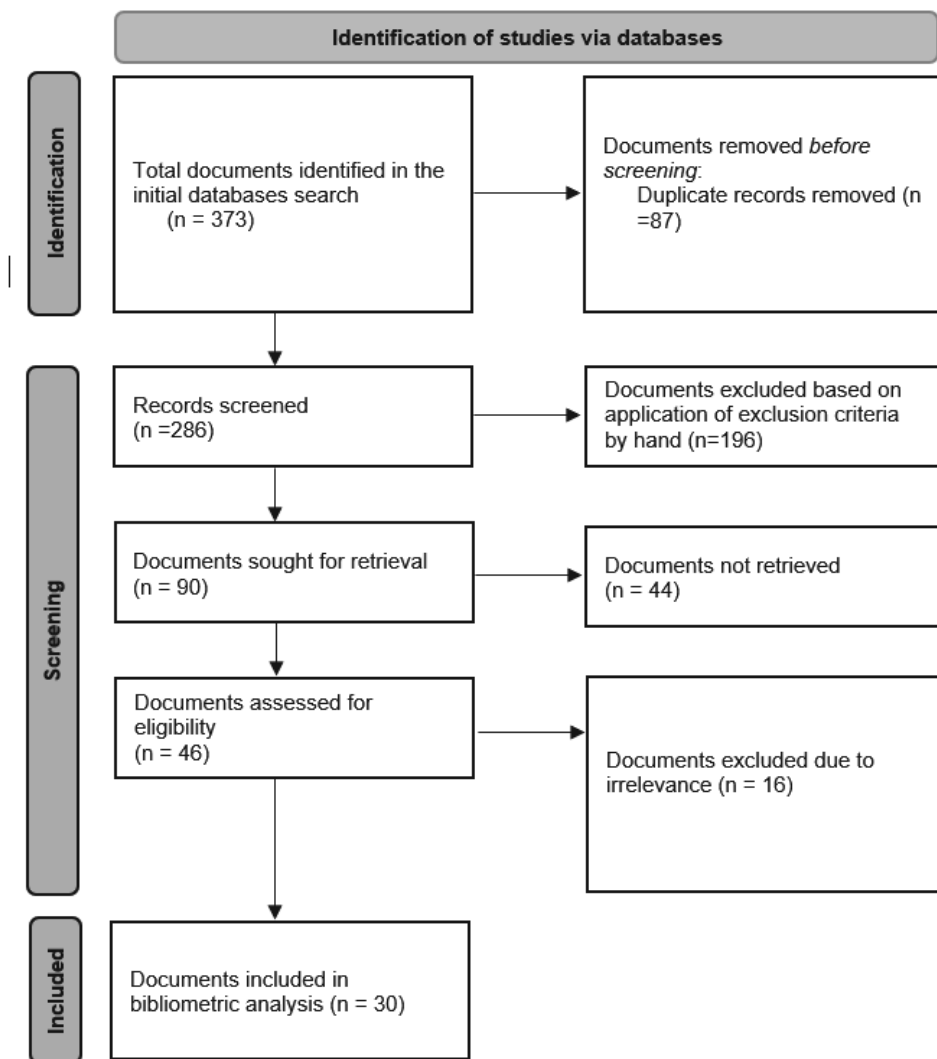
Source: Authors' research

Although the research was initially intended to focus exclusively on the terms forensic accounting and non-financial or sustainability reporting, the fact that the database searches yielded only a very small number of papers led to the need to introduce the concepts of fraud, manipulation, and greenwashing, which resulted in a larger number of initially potentially relevant papers.

4. Research results

In the search for previous studies on the application of forensic accounting in detecting potential fraud in sustainability reporting, bibliometric analysis was used. Bibliometric analysis uses quantitative methods to examine patterns of scientific publication. The results of bibliometric analysis serve to support the development of the theoretical framework and the presentation of previously conducted research.

Before conducting the bibliometric analysis itself, the PRISMA method was implemented. The PRISMA method was used for collecting information in order to determine whether it was possible to conduct a bibliometric analysis and identify scientific articles (Page et al., 2021). In this research, a comprehensive screening process was conducted.

Figure 1: PRISMA flow diagram

Source: Authors' research

The PRISMA flow diagram in Figure 1 illustrates the main procedures used for the systematic selection of publications for the bibliometric study. By searching four databases—WoS, Scopus, ProQuest, and The Lens—a total of 539 records were initially identified. Since the same records appeared across multiple search strings within the same database, the records were first deduplicated within each database. After consolidation, 373 records entered the identification stage. Following the removal of duplicates across the databases, 286 records remained. Their titles,

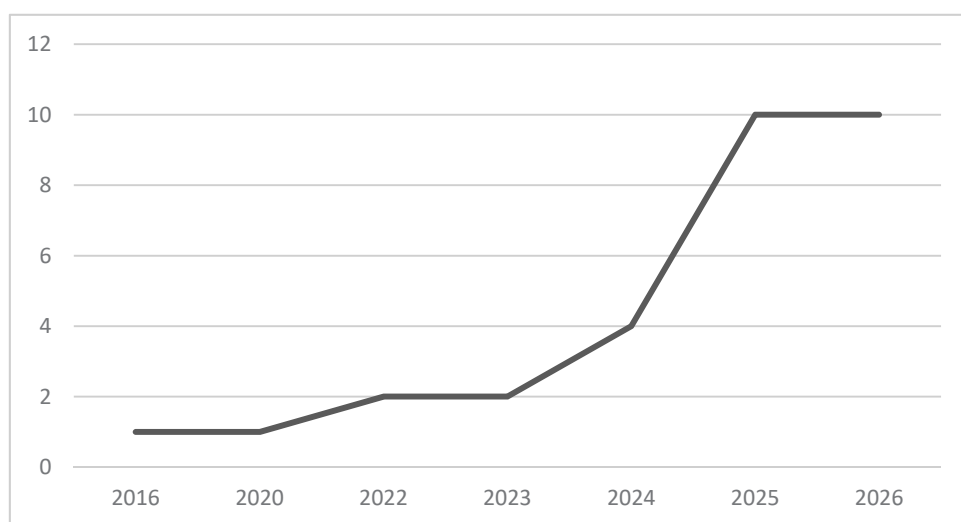
abstracts, and keywords were screened in order to determine their relevance to the research topic.

After the screening process, 196 records were excluded, leaving 90 papers. Full texts were then sought for these 90 papers. However, for some papers, the full text was not available or the paper was not written in English. As a result, an additional 44 papers were excluded, leaving 46 papers for further review. Following a more detailed assessment, a further 16 papers were considered not sufficiently relevant to the research topic. Consequently, a final sample of 30 papers was included in the bibliometric analysis.

Papers dealing exclusively with fraud in financial statements were excluded from the bibliometric analysis, as were papers addressing forensic accounting without any sustainability dimension, studies focusing exclusively on ESG performance without examining the credibility and reliability of disclosures, conference abstracts, book reviews, and other non-research publications. The final bibliometric sample therefore included papers examining forensic accounting or forensic auditing in connection with sustainability, ESG, or non-financial reporting, as well as studies investigating fraud and manipulation in sustainability reporting or non-financial information.

The papers included in the bibliometric analysis were published between 2016 and 2026. As shown in Graph 1, an increasing publication trend can be observed over this ten-year period. This indicates a growing academic interest in sustainability reporting, as well as in the detection of potential fraud related to non-financial and sustainability information.

Graph 1: Publication trend over the years (productivity)

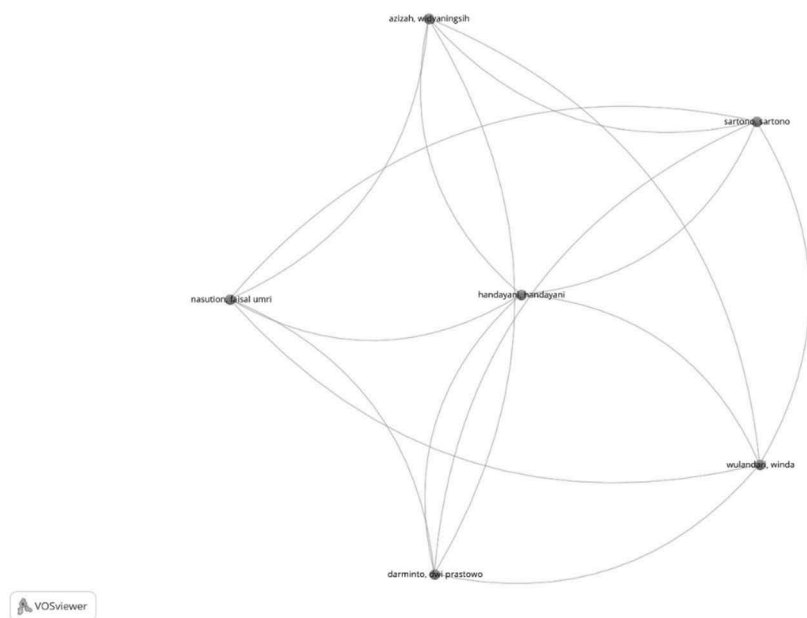


Source: Authors' research

An analysis of the most productive journals showed that no journal had published more than one paper on the topic of sustainability and forensic accounting or fraud detection. The papers included in the bibliometric analysis were published across 30 different publication outlets. This indicates that there is no clear concentration of research within one or several specialized journals. Therefore, it may be concluded that this interdisciplinary field remains underexplored, although research interest has increased in recent years. Among the identified journals, the Sustainability Accounting, Management and Policy Journal is particularly relevant to the research topic due to its focus on accounting, sustainability, management, and related reporting practices. Specialized journals such as the Journal of Auditing, Finance and Forensic Accounting and the Journal of Forensic Accounting Research are also represented.

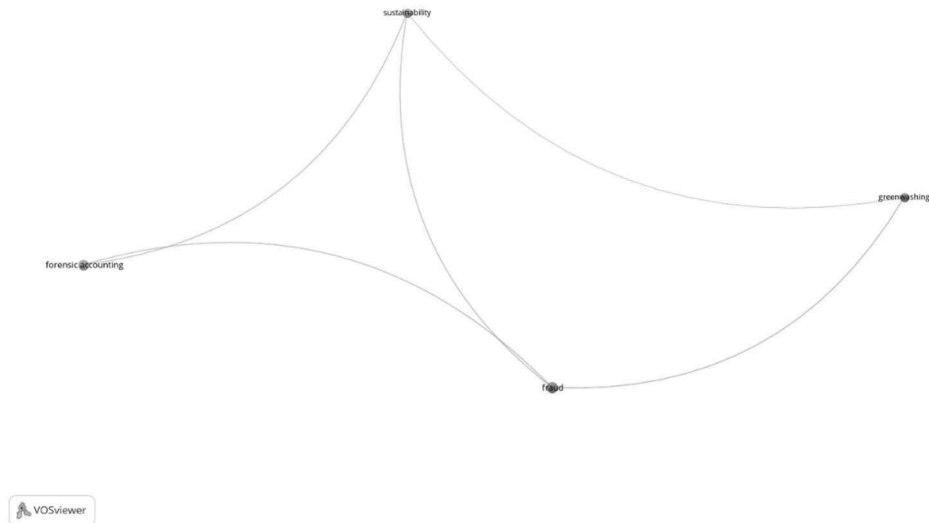
The VOSviewer software was used to present the data collected through the bibliometric analysis. This software enables the creation of maps of authors and journals based on citation data, as well as keyword maps based on their occurrence (Van Eck and Waltman, 2010).

A co-authorship analysis was conducted in VOSviewer to visualize collaboration among researchers based on jointly published scientific papers, as shown in Figure 2. The map displays only six authors who form an interconnected network. All of them belong to a single cluster, meaning that no separate research groups were identified. Although the sample contains 30 papers and a considerably larger number of authors, only this small group met the criteria for inclusion in the connected co-authorship network. This suggests that author collaboration within the selected papers included in the analysis is very limited and fragmented. These results suggest that research on the relationship between forensic accounting and sustainability reporting has not yet developed a stable and interconnected research community.

Figure 2: Co-authorship analysis

Source: Authors' research

A co-occurrence analysis was then conducted in VOSviewer in order to examine the content and concepts within the scientific papers, that is, to identify keywords that frequently appear together in the same papers and thereby enable the mapping of thematic structures and trends. The co-occurrence analysis is presented in Figure 3. The map shows only four dominant interconnected concepts: forensic accounting, sustainability, fraud, and greenwashing, which are divided into two clusters. The green cluster consists of forensic accounting and sustainability, indicating that there is a recognizable conceptual relationship between these two concepts in the analysed papers.

Figure 3: Co-occurrence analysis

Source: Authors' research

The second cluster (red) consists of the combination of fraud and greenwashing. Fraud represents a traditional area of interest in forensic accounting, while greenwashing represents one of the potential problems related to the credibility of sustainability-related information. The map shows that the term fraud is positioned between forensic accounting/sustainability and greenwashing and is linked to the other concepts. Based on the analysed papers, it may be concluded that fraud can represent a conceptual bridge between traditional forensic accounting issues and emerging issues related to the credibility of sustainability information. The map also shows that greenwashing is linked to sustainability and fraud, while there is no direct link between forensic accounting and greenwashing. This may indicate that the potential application of forensic accounting for identifying greenwashing has not yet been sufficiently addressed in the analysed literature. In the sample of 30 papers, only a small number of keywords reached the threshold required for inclusion in the co-occurrence network. This may suggest that there is still a limited body of literature at the intersection of these fields. Furthermore, authors may use different terms for similar concepts, and there may be a lack of a unified conceptual framework because the field is still at a relatively early stage of development. Both the co-authorship analysis and the keyword co-occurrence analysis indicate a research field that remains relatively fragmented and insufficiently consolidated.

Among the papers included in the bibliometric analysis, a particularly small group of studies explicitly addressed the intersection between forensic accounting/auditing

and sustainability or ESG reporting. These studies differ in their methodological approaches and in how directly they address this relationship. Azizah et al. (2025) conduct a systematic review of the role of forensic auditing and digital forensics in the integrity of sustainability disclosures, with the aim of detecting fraud, verifying the authenticity of non-financial data, and preventing greenwashing. Hossain et al. (2025) also examine forensic auditing in fraud detection and in ensuring compliance with ESG reporting standards, including greenwashing and social-compliance misreporting.

Sun (2025) provides a specific forensic/technology framework. Sun argues that the problem of ESG reporting is not limited to greenwashing in a communication-related sense, but also includes manipulation of specific ESG data. Typical forms of misconduct include selective disclosure, data distortion, exaggeration of sustainability claims, lack of transparency in supply chains, and misuse of or excessive reliance on third-party certifications. The author particularly emphasizes the possibility that a company may formally appear highly sustainable, while its actual operations or supply chain do not support what has been disclosed. An important finding is that Sun considers traditional auditing insufficient for some more complex forms of ESG fraud. The author identifies fragmented standards, technological limitations, reliance on self-reported ESG data, and possible conflicts of interest as key problems.

It is precisely in this context that forensic accounting is introduced as an extended investigative approach that does not merely ask whether information has been formally presented correctly, but also whether it can be demonstrated that the reported claim corresponds to the actual situation. The author sees the combination of financial auditing, legal analysis, and digital forensics as a key feature distinguishing forensic accounting from a conventional audit approach. For example, if a company claims that it is reducing emissions, a forensic approach does not necessarily have to rely solely on the company's internal ESG dataset but may compare it with other sources, supply-chain data, or geospatial evidence. Sun therefore proposes a multidimensional verification framework combining blockchain traceability, AI-based semantic analysis, and satellite monitoring.

Finata et al. (2026) view forensic accounting as an independent monitoring mechanism used to test the relationship between information presented in a sustainability report and actual operational data. According to the authors, the role of the forensic accountant is particularly important because, in addition to determining whether a company has disclosed certain information, it is also necessary to verify whether evidence exists to support the claim; whether the disclosed information is materially significant; whether the narrative corresponds to the magnitude of the actual impact; whether the trend is consistent over several years; and whether there are differences between ESG reports and operational data. Finata et al. (2026) also examine the role of artificial intelligence in their study, arguing that the forensic accountant provides the investigative concept, professional judgement,

and evidence verification, while AI enables the analysis of large volumes of sustainability reports.

Fauziah et al. (2026) empirically demonstrate the extent to which forensic accounting is weakly integrated into the existing sustainability/greenwashing literature. The authors also conducted a bibliometric analysis, and the results showed that forensic accounting does not constitute a separate dominant research stream. The authors conclude that previous research has extensively examined why companies engage in greenwashing and what its consequences are, while considerably less attention has been devoted to how it can be investigatively demonstrated that a specific ESG claim has been manipulated.

Conclusion

The aim of this paper was to determine whether there are studies in the field of forensic accounting and sustainability reporting. Four databases were used in the search for relevant papers: WoS, Scopus, ProQuest, and The Lens. In the two main databases, WoS and Scopus, the exact query “forensic accounting” AND “sustainability reporting” returned no records. Searches in the additional databases produced broader sets of records, but subsequent screening showed that only a limited number of papers explicitly addressed the direct intersection of the two fields. The expanded search showed that relevant research appears through research streams focused on fraud, information manipulation, greenwashing, ESG reporting, the credibility of non-financial information, and forensic auditing.

The results of the bibliometric analysis of the final sample of 30 publications confirmed the fragmentation of the research field. The keyword co-occurrence analysis revealed a relatively sparse conceptual network in which forensic accounting, sustainability, fraud, and greenwashing emerged as the central concepts. Particularly significant is the role of fraud, which appears as a conceptual link between the traditional field of forensic accounting and issues related to sustainability and greenwashing. At the same time, the limited co-authorship network indicates that a broader and stable research community systematically linking forensic accounting with sustainability reporting has not yet developed. The results support the conclusion that this is a new and still insufficiently consolidated research field.

The analysed papers indicate that sustainability reporting (ESG reporting) is increasingly facing the same problems that led to the development of forensic accounting in financial reporting: data manipulation, selective disclosure, concealment of unfavourable information, fraud risk, greenwashing, and problems related to data credibility. Therefore, the role of forensic accounting could be extended from traditional financial reporting to sustainability reporting. Just as forensic accounting has both a detective and preventive function in the analysis of financial statements, it could also play an important role in the analysis of

sustainability reports, which contain a large number of quantitative and qualitative indicators, thereby contributing to the strengthening of corporate governance.

This study has several limitations that should be taken into account when interpreting the results. The literature search was conducted across four databases, which means that some relevant studies indexed in other sources may not have been included in the analysis. Furthermore, the results of the systematic search depend on the selected combination of keywords. Since the field still does not have fully standardized terminology, some papers may examine similar issues without explicitly using the term forensic accounting. A second limitation relates to the relatively small final sample of 30 publications. Although, on the one hand, the small number of relevant papers indicates the limited development of the field, it also limits the possibility of presenting more extensive bibliometric networks and generalizing the results. Finally, the existing literature still does not provide sufficient empirical evidence to conclude that the application of forensic accounting can improve the quality of sustainability reporting or reduce greenwashing. Most of the recent literature identifies potential applications and proposes relevant techniques, but their effectiveness has not yet been sufficiently tested in real organizational settings. Therefore, the results of this study should be interpreted as confirmation of significant research potential rather than as evidence of the already established effectiveness of forensic accounting in sustainability reporting.

Future research should empirically examine the possibility of applying specific forensic accounting techniques using data from sustainability and ESG reports. Future studies should also consider the development of integrated models for the forensic verification of financial and non-financial information. Since ESG activities may have financial consequences, while financial decisions may at the same time be linked to disclosed sustainability indicators, the separation between financial fraud and non-financial reporting fraud is becoming increasingly difficult to justify. The development of a hybrid assurance approach could enable the simultaneous verification of financial data, ESG metrics, and the evidence supporting sustainability claims.

The contribution of this paper lies in the systematization of the still dispersed literature at the intersection of forensic accounting and sustainability reporting, as well as in the identification of conceptual links between forensic accounting, fraud, sustainability, and greenwashing. The results show that the issue of the credibility of non-financial information is increasingly approaching the traditional area of interest of forensic accounting, while the direct theoretical and, particularly, empirical integration of these two fields remains insufficiently developed. This creates space for the development of a new research direction focused on the application of forensic methods in verifying the integrity of sustainability and ESG information. As corporate reporting expands from financial information to sustainability and ESG information, the issues of fraud and manipulation can no longer be limited exclusively to financial statements; consequently, it is necessary to consider

extending the traditional role of forensic accounting to the verification of the credibility of non-financial reporting.

References

- Azizah, W., Handayani, H., Sartono, S., Nasution, F. U., Wulandari, W., & Darminto, D. P. (2025). Forensic Audit: The Integrity Backbone of Sustainability Disclosure. *International Journal of Social Science and Human Research*, 8(10). <https://doi.org/10.47191/ijsshr/v8-i10-65>
- Bakhit, M. A. A. (2024). The role of forensic accounting to reduce financial corruption practices and improving the quality of financial reports. *Revista de Gestão Social e Ambiental*, 18(3), e04931. <https://doi.org/10.24857/rgsa.v18n3-090>
- Budimir, N. (2017). Forenzično računovodstvo i forenzična revizija. *Elektronički zbornik radova Veleučilišta u Šibeniku*, 11 (1-2), 101-113.
- DiGabriele, J. A., & Huber, W. D. (2015). Topics and methods in forensic accounting research. *Accounting Research Journal*, 28(1), 98–114. <https://doi.org/10.1108/ARJ-08-2014-0071>
- Fauziah, D. A., Syafii, S., Abidin, K., & Hermanto, B. A. (2026). Greenwashing, ESG Fraud, and the Forensic Gap: A Bibliometric Mapping. *MARGIN ECO*, 10(1), 251–271. <https://doi.org/10.32764/margineco.v10i1.7319>
- Finata, E. A., Azarina, A., Lisnawati, M., & Rahayu, H. A. (2026). THE ROLE OF FORENSIC ACCOUNTING AND ARTIFICIAL INTELLIGENCE TO UNCOVERING GREENWASHING PRACTICES IN SUSTAINABILITY REPORTS. *Jurnal Akuntansi Dan Ekonomi Bisnis*, 15(1), 87–95. <https://doi.org/10.33795/jaeb.v15i01.8814>
- Harzing, A. (2014). A longitudinal study of Google Scholar coverage between 2012 and 2013. *Scientometrics*, 98(1), 565-575. <https://doi.org/10.1007/s11192-013-0975-y>
- Harzing, A. & Alakangas, S. (2016). Google Scholar, Scopus and Web of Science: a longitudinal and cross-disciplinary comparison. *Scientometrics*, 106, 787-804. <https://doi.org/10.1007/s11192-015-1798-9>
- Hossain, M. Z., Hredoy, A. I., & Rashed, J. R. (2025). Forensic Auditing and ESG Compliance: Detecting Fraud and Enhancing Transparency in Corporate Reporting. *European Journal of Management, Economics and Business*, 2(6), 108–123. [https://doi.org/10.59324/ejmeb.2025.2\(6\).08](https://doi.org/10.59324/ejmeb.2025.2(6).08)
- Monserrat, C. (2026). Bibliometric analysis of big data applications in accounting fraud detection. *Intangible Capital*, 22(1), 101–122. <https://doi.org/10.3926/ic.3119>
- Okeke, D. F. C. S. A., Ekerare, D. N. J., Obidike, D. P. C., & Okwu, P. I. (2022). Forensic Accounting and Integrated Financial Reporting of Banks using

- Hausman. *International Journal of Advanced Engineering, Management and Science*, 8(6), 26–36. <https://doi.org/10.22161/ijaems.86.4>
- Ozili, P. K. (2025). Forensic accounting research around the world. *Journal of Financial Reporting and Accounting*, 23(1), 128–153. <https://doi.org/10.1108/JFRA-02-2023-0106>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, 88, 105906. <https://doi.org/10.1016/j.ijssu.2021.105906>
- Poiriazi, E., Zournatzidou, G., Konteos, G., & Sariannidis, N. (2025). Analyzing the interconnection between environmental, social, and governance (ESG) criteria and corporate corruption: Revealing the significant impact of greenwashing. *Administrative Sciences*, 15(3), 100. <https://doi.org/10.3390/admsci15030100>
- Rezaee, Z. (2019). *Forensic Accounting and Financial Statement Fraud*, Volume II: Forensic Accounting Performance. Business Expert Press
- Sun, Q. (2025). Forensic Accounting Approaches to Detecting and Preventing Greenwashing and Data Manipulation in ESG Reporting. *Advances in Economics, Management and Political Sciences*, 199(1), 160–166. <https://doi.org/10.54254/2754-1169/2025.bj25284>
- Tarjo, T., Anggono, A., Said, J., Sakti, E. (2024). Environmental, Social, and Governance (ESG) Fraud: A Bibliometric Study and Systematic Literature Review. *E3S Web of Conferences*, 499, 1-13. <https://doi.org/10.1051/e3sconf/202449901002>
- Van Eck, N. J., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523–538. <https://doi.org/10.1007/s11192-009-0146-3>

Appendix: Documents included in bibliometric analysis

- Aldeikeh, N. A., Alshurafat, H., Ananzeh, H., & Alzoubi, A. B. (2026). Forensic accounting as a tool for sustainable public institutions: Evidence from money laundering prevention in Jordan. *Discover Sustainability*, 7(1), 614. <https://doi.org/10.1007/s43621-026-02963-x>
- Anggraeni, A. A., & Dewi, H. P. (2022). Green Accounting and Corporate Social Responsibility Disclosure: Financial Performance of Mining Companies in Indonesia. *Journal of Auditing, Finance, and Forensic Accounting*, 10(1), 61–74. <https://doi.org/10.21107/jaffa.v10i1.14034>
- Azizah, W., Handayani, H., Sartono, S., Nasution, F. U., Wulandari, W., & Darminto, D. P. (2025). Forensic Audit: The Integrity Backbone of Sustainability Disclosure. *International Journal of Social Science and Human Research*, 8(10). <https://doi.org/10.47191/ijsshr/v8-i10-65>
- Cinceoglu, V., Strauss, N., Binder-Tietz, S., & Alvarado, I. O. (2026). Addressing Greenwashing-The Role of Strategic Communication in ESG Reporting. *International Journal of Strategic Communication*, 20(2, SI), 173–194. <https://doi.org/10.1080/1553118X.2025.2591309>
- Farnsel, C., Preber, B., Ramamoorti, S., & Venkatachalam, N. (2025). ESG Fraud Risks: A Research Agenda. *Journal of Forensic Accounting Research*, 1–25. <https://doi.org/10.2308/jfar-2024-021>
- Fauziah, D. A., Syafii, S., Abidin, K., & Hermanto, B. A. (2026). Greenwashing, ESG Fraud, and the Forensic Gap: A Bibliometric Mapping. *MARGIN ECO*, 10(1), 251–271. <https://doi.org/10.32764/margineco.v10i1.7319>
- Febrina, I., & Kuraesin, A. D. (2025). The Role of Corporate Reputation in Moderating the Effects of Greenwashing, Fraud, and SDGs Disclosure on Firm Value. *Architecture Image Studies*, 6(3), 839–850. Scopus. <https://doi.org/10.62754/ais.v6i3.334>
- Finata, E. A., Azarina, A., Lisnawati, M., & Rahayu, H. A. (2026). The Role of Forensic Accounting and Artificial Intelligence to Uncovering Greenwashing Practices in Sustainability Reports. *Jurnal Akuntansi Dan Ekonomi Bisnis*, 15(1), 87–95. <https://doi.org/10.33795/jaeb.v15i01.8814>
- Furtado, N. G., Furtado, J. V., & Filgueiras, J. R. (2026). Blockchain for Consumer Trust and Engagement in Sustainable Forestry: A Permissioned Ledger Approach to ESG Audits. Driving Engagement and Loyalty Through Customer-Centric Digital Marketing, 105–151. Scopus. <https://doi.org/10.4018/979-8-3373-4662-5.ch004>
- Gupta, M., Aggarwal, P. K., & Gupta, R. (2023). Corporate Sustainability with Forensic Auditing: An Evidence from India. *International Journal of Experimental Research and Review*, 35, 128–137. Scopus. <https://doi.org/10.52756/ijerr.2023.v35spl.012>

- Harjoto, M. A., & Semba, H. D. (2026). Sustainability and financial disclosure: Role of ESG in key audit matters adoption. *Managerial Auditing Journal*, 41(1), 31. ProQuest One Business (3289770164). <https://doi.org/10.1108/MAJ-01-2025-4663>
- Hidayah, R., & Ratmono, D. (2025). The Impact of Audit Committee on Greenwashing: Forensic Science in Managerial Sustainability Governance. IOP Conference Series. Earth and Environmental Science, 1499(1), 012034. Publicly Available Content Database (3216356318). <https://doi.org/10.1088/1755-1315/1499/1/012034>
- Hossain, M. Z., Hredoy, A. I., & Rashed, J. R. (2025). Forensic Auditing and ESG Compliance: Detecting Fraud and Enhancing Transparency in Corporate Reporting. European Journal of Management, *Economics and Business*, 2(6), 108–123. [https://doi.org/10.59324/ejmeb.2025.2\(6\).08](https://doi.org/10.59324/ejmeb.2025.2(6).08)
- Kazemi, A., Mehrani, S., & Homayoun, S. (2025). Risk in Sustainability Reporting: Designing a DEMATEL-Based Model for Enhanced Transparency and Accountability. *Sustainability*, 17(2). <https://doi.org/10.3390/su17020549>
- Kurpierz, J. R., & Smith, K. (2020). The greenwashing triangle: Adapting tools from fraud to improve CSR reporting. *Sustainability Accounting Management and Policy Journal*, 11(6, SI), 1075–1093. <https://doi.org/10.1108/SAMPJ-10-2018-0272>
- Lucut Capras, I., Achim, M. V., Rahat, B., & Nicolo, G. (2025). Sustainable auditing governance in a changing landscape: Cross-country evidence on the effect of traditional and innovative audit mechanisms on fraud risk. *Journal of Innovation & Knowledge*, 10(5). <https://doi.org/10.1016/j.jik.2025.100798>
- Nagnonhou, S. R. B., Imoniana, J. O., Reginato, L., & Silva, W. L. (2023). Role of Connectors in Corporate Fraud and Corruptions in Era of Circular Economy. *Social Sciences*, 12(3), 134. <https://doi.org/10.3390/socsci12030134>
- Natanson, N., Lai, H., Bothé, S., & Florén, J. (2026). Can the Finance Sector Be Trusted?: Greenwashing, Greenhushing and Greenwishing. *Key Debates in Sustainable Investment*, 165–185. Routledge. <https://doi.org/10.4324/9781003467632-11>
- Okeke, D. F. C. S. A., Ekerare, D. N. J., Obidike, D. P. C., & Okwu, P. I. (2022). Forensic Accounting and Integrated Financial Reporting of Banks using Hausman. International Journal of Advanced Engineering, *Management and Science*, 8(6), 26–36. <https://doi.org/10.22161/ijaems.86.4>
- Okmeydan, S. B. (2025). A Colorful Lie on Ecological and Environmental Sustainability: Greenwashing. The Different Aspects of Greenwashing, 1–41. Scopus.
- Pham, Q. H., & Vu, K. P. (2024a). Insight into how cyber forensic accounting enhances the integrated reporting quality in small and medium enterprises. *Cogent Business & Management*, 11(1). Coronavirus Research Database; ProQuest One

Business; Publicly Available Content Database (3152074360). <https://doi.org/10.1080/23311975.2024.2364053>

Pham, Q. H., & Vu, K. P. (2024b). Optimization of Accounting information System in Public Sector for Sustainable Risk Management Under Big Data Analytics. Does forensic Accountants' Skill Generate Differences? *Foundations of Management*, 16(1), 67–82. Coronavirus Research Database; ProQuest One Business; Publicly Available Content Database (3063265784). <https://doi.org/10.2478/fman-2024-0005>

Shivdas, A. (2025). The Role of Ethics and Sustainability Reporting (BRSR) in Reducing Financial Frauds in Indian Companies. *Quest Journal of Management*, 16(1), 17–26. ProQuest One Business (3263167634).

Singh, H., & Singh, A. (2024). Blockchain and ESG: Unfolding Landscape of Sustainable Reporting. *Sustainability Reporting and Blockchain Technology*, 3–14. Scopus. <https://doi.org/10.4324/9781003378341-2>

Steinmeier, M. (2016). Fraud in Sustainability Departments? An Exploratory Study: JBE. *Journal of Business Ethics*, 138(3), 477–492. ProQuest One Business (1826089720). <https://doi.org/10.1007/s10551-015-2615-3>

Sun, Q. (2025). Forensic Accounting Approaches to Detecting and Preventing Greenwashing and Data Manipulation in ESG Reporting. *Advances in Economics, Management and Political Sciences*, 199(1), 160–166. <https://doi.org/10.54254/2754-1169/2025.bj25284>

Sunde, T. V., & Olsen, C. (2026). The Problem of Data Integrity in Sustainability Accounting and Assurance: Blockchain, Artificial Intelligence (AI), and Triple-Entry Accounting (TEA). *Artificial Intelligence in Sustainability Accounting and Assurance*, 175–197. Scopus. <https://doi.org/10.4324/9781003615668-10>

Sutrisno, S., Djaddang, S., & Suratno, S. (2026). Antecedents of Pentagon Fraud against Sustainability Reporting and Behavioral Organization as Intervening. *International Review of Management and Marketing*, 16(3), 452–465. Scopus. <https://doi.org/10.32479/irmm.22872>

Tan, Y., Qi, L., Chen, G., Ngan, I. H., & Wu, L. (2026). Mandatory ESG disclosure and greenwashing: Evidence from a quasi-natural experiment in China. *Journal of Environmental Planning and Management*, 1–28. <https://doi.org/10.1080/09640568.2026.2633736>

Xanthopoulou, A., Skordoulis, M., Kalantonis, P., & Arsenos, P. (2024). Integrating Corporate Governance and Forensic Accounting: A Sustainable Corporate Strategy Against Fraud. *Journal of Governance and Regulation*, 13(2 Special Issue), 327–338. Scopus. <https://doi.org/10.22495/jgrv13i2siart9>

Dejan Jakšić
University of Novi Sad, Faculty of
Economics in Subotica
ORCID: 0000-0003-1711-1869

Kristina Peštović
University of Novi Sad, Faculty of
Economics in Subotica
ORCID: 0000-0001-9851-1025

FORENSIC PROCESS MINING: A PYTHON-BASED APPROACH TO DETECTING RED FLAGS IN BUSINESS EVENT LOGS

UDK 004.9:657.6

DOI <https://doi.org/10.54318/czf.2026.dj.011>

Abstract

Digitalization of business processes has shifted a significant part of forensic accounting evidence from paper-based documentation to information system traces. Enterprise resource planning, workflow, banking and other systems record events that indicate which activity was performed, when, by whom and within which business case. This paper conceptualizes forensic process mining as the application of process mining techniques, forensic accounting logic and audit analytics to business event logs in order to detect red flags related to irregular process behaviour, control violations, anomalous timing, unusual actor involvement and financially relevant deviations. The paper combines a review of contemporary literature on process mining in auditing and fraud-related analytics with a proof-of-concept case study based on a simulated procurement process event log. A Python prototype, built on PM4Py and other open-source libraries, was developed to generate an interactive forensic-oriented process mining report. The prototype analyses process variants, directly-follows relations, Petri-net-based conformance, bottlenecks, service times, resource workload, handover-of-work patterns, temporal heatmaps, service-level agreement breaches, Benford indicators and z-score outliers. The results show that a customizable Python approach can transform event log data into structured red flags for further audit or forensic investigation. ChatGPT was used to assist in generating the simulated event log and in selected code-drafting, debugging and refinement tasks. The authors subsequently reviewed the analytical outputs and retained responsibility for the methodological choices and interpretation of the results. However, detected red flags do not represent evidence of fraud; they are risk indicators that require professional judgment, documentation review and additional forensic procedures.

Key words: forensic process mining; forensic accounting; event logs; red flags; Python; PM4Py; fraud risk

JEL Classification: M42, M41, C88

Introduction

Business activities are increasingly executed through digital systems that continuously record operational events. Enterprise resource planning systems, workflow applications, e-procurement platforms, banking systems and customer relationship management systems do not only store financial transactions; they also store detailed traces of the processes through which those transactions are initiated, approved, modified, delivered, posted and paid. In this context, forensic accounting evidence is no longer limited to invoices, contracts, journal entries and supporting documents. It also includes event logs that show the sequence of activities, the timing of actions, the actors involved and the contextual attributes of individual business cases.

This shift is important for forensic accounting and auditing. Traditional audit and forensic procedures often focus on documents, accounting records and transaction-level anomalies. However, irregularities in a digital environment may also appear as process-level patterns: skipped approvals, unusual sequences of activities, repeated approval steps, manual corrections without documentation, unusually short control execution, long waiting times, concentration of critical actions within one user, or atypical process paths for specific suppliers. Such patterns may not be visible if the analysis is limited to accounting entries or static reports.

Process mining provides a suitable analytical foundation for this problem. It uses event data to discover actual processes, compare observed behaviour with expected models and analyse performance and compliance issues. Werner, Wiese, and Maas (2021) argue that process mining can support financial statement audits by analysing business processes and internal controls more effectively in highly digitalized transaction-processing environments. Similarly, Jans and Eulerich (2022) position process mining as a relevant technique for financial auditing because it enables auditors to analyse business processes using data generated by information systems.

The objective of this paper is to conceptualize forensic process mining as a targeted application of process mining in forensic accounting and to demonstrate, through a Python-based proof of concept, how business event logs can be analysed to detect red flags relevant for fraud risk assessment and further forensic investigation. Generative artificial intelligence (AI) is not the central topic of the paper, but it is briefly considered as a practical enabler that may help forensic accountants translate red flag logic into executable analytical tools.

The paper proceeds by discussing process mining and event logs, defining forensic process mining, presenting the Python prototype, demonstrating the procurement case study, and discussing contributions and limitations.

1. Process Mining and Business Event Logs in Audit Context

Process mining is commonly positioned at the intersection of data science and process science. According to van der Aalst (2022), process mining uses event data

and process models to provide insights, identify bottlenecks and deviations, diagnose compliance and performance problems, and support process improvement. Unlike traditional business intelligence, which often reduces behaviour to aggregated indicators, process mining analyses end-to-end process execution at the level of individual events.

An event log is the basic data input for process mining. Each event usually contains at least a case identifier, an activity name and a timestamp. Optional attributes may include resource, location, cost, supplier, customer, amount, document type or organizational unit (van der Aalst, 2022). In accounting and auditing, these optional attributes are particularly important because they connect the control-flow perspective with financial and organizational context. A procurement event log, for example, may show not only that an invoice was posted after goods receipt, but also which employee posted it, which supplier was involved, what amount was recorded and how much time elapsed between the previous and subsequent activities.

Three process mining disciplines are especially relevant for audit and forensic work: process discovery, conformance checking and process enhancement. Process discovery creates a model of the actual process from event logs. Conformance checking compares observed behaviour with a discovered or normative process model. Process enhancement adds information such as frequencies, waiting times, service times, costs and resource involvement to the process model (Berti et al., 2023; van der Aalst, 2022). These techniques can help auditors move from interview-based and sample-based understanding of processes to a more data-driven assessment of what actually happened in the information system.

The audit literature increasingly recognizes this potential. Werner et al. (2021) explain that traditional audit procedures become inefficient in environments characterized by integrated information systems and large volumes of automated or semi-automated transactions. Process mining can analyse recorded business transactions, identify deviations automatically and guide further substantive audit procedures. Föhr et al. (2025) add that process mining can enhance the evaluation of internal control systems and tests of controls through comprehensive analysis of process variants. Their audit process Mining framework includes phases such as business process understanding, event data understanding, event log creation, process mining analysis and audit objective evaluation.

Process mining is also relevant beyond external financial statement audits. Eulerich et al. (2025) link process mining to the three-lines model and argue that operational management, risk and compliance functions, and internal audit can all use process mining to identify risks, control violations and anomalous activities. This is highly relevant for forensic accounting because fraud prevention and detection often require cooperation across operational controls, compliance, risk management and independent assurance functions.

Nevertheless, process mining is not automatically forensic. Standard process mining analyses usually aim to understand, improve and monitor processes. Forensic

accounting requires a more specific orientation: identifying indicators that may point to irregular behaviour, control weaknesses, potential manipulation or fraud risk. Therefore, the process mining perspective must be combined with forensic accounting logic, audit analytics and domain-specific red flag rules.

2. From Process Mining to Forensic Process Mining

Forensic process mining can be defined as the application of process mining techniques, forensic accounting logic and audit analytics to business event logs with the aim of identifying red flags related to irregular process behaviour, control violations, anomalous timing, unusual actor involvement and financially relevant deviations. This definition emphasizes that forensic process mining is not merely a visualization of process flows. It is a targeted analytical approach that transforms event log data into indicators for further forensic inquiry.

The distinction between a red flag and evidence is essential. A red flag is not proof of fraud. It is an indicator that a case, activity, user, supplier, time pattern or transaction deserves further professional attention. Forensic conclusions require additional evidence, such as source documents, contracts, system authorizations, interviews, accounting records and legal assessment. This is consistent with the fraud audit perspective of Rodríguez-Quintero et al. (2021), who emphasize the role of process mining and visual analytics in identifying relevant fraud-related patterns during audits, rather than replacing professional investigation.

Forensic process mining can use several categories of red flags, as shown in Table 1.

Table 1: Categories of forensic red flags in business event logs

Category	Examples of red flags	Forensic relevance
Control-flow red flags	skipped approval, repeated approval, procedure bypass, manual correction without documentation	possible violation of internal procedures or intentional circumvention of controls
Time-based red flags	SLA breach, unusually long waiting time, unusually short execution of control activities	possible bottleneck, ineffective control, backdating or process manipulation
Resource-based red flags	same person creates and approves a transaction, unusual workload concentration, atypical handover	segregation of duties risk or unusual actor involvement

Category	Examples of red flags	Forensic relevance
Case- or supplier-based red flags	specific supplier follows rare variants, high-value cases follow shortened path	possible preferential treatment, collusion risk or targeted process manipulation
Financial red flags	outlier amounts, unusual first-digit distribution, amounts close to authorization thresholds	possible manipulation of transaction values or threshold avoidance
Conformance red flags	log-only activities, model-only activities, deviation from expected path	mismatch between expected and observed process behaviour
Behavioural red flags	activities outside usual working hours, repeated activity loops, rare user-activity combinations	unusual user behaviour requiring further explanation

Source: Authors' proposal.

This classification extends the traditional focus of forensic accounting. Instead of asking only whether a transaction amount is unusual, forensic process mining asks whether the way in which the transaction moved through the process is unusual. A procurement case may bypass evaluation, be approved unusually quickly, be handled by the same person in incompatible roles, or follow a rare path linked to one supplier.

The literature supports this orientation. In financial statement audits, process mining can help identify uncommon control-flow paths, process variants and manual postings not covered by the mined model (Föhr et al., 2025; Werner et al., 2021). Fraud-related studies also show applications in visual fraud audit, forensic event log investigation and digital onboarding fraud detection (Khan et al., 2023; Rodríguez-Quintero et al., 2021; da Silva et al., 2023).

However, standard process mining software is not necessarily designed as a forensic accounting tool. Process mining tools are valuable for discovery, conformance checking, performance analysis and dashboards, but forensic work often requires customized rules: segregation of duties tests, supplier-specific anomaly analysis, threshold-based authorization testing, exception reporting, case-level risk scoring and integration of financial attributes. Drakoulogkonas and Apostolou (2021) show that process mining tools can be compared using multiple criteria and that tool selection depends on user needs. In a forensic setting, this implies that graphical tools may be useful, but programmable approaches become important when the auditor needs to translate specific red flag logic into executable tests.

3. Python-Based Development of a Forensic Process Mining Prototype

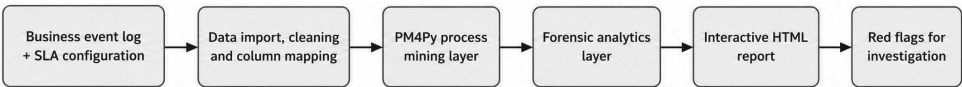
The proof-of-concept prototype presented in this paper was developed in Python. Python was selected because it provides flexibility, open-source libraries, transparent code, integration with data analysis tools and the possibility of implementing customized forensic rules. The prototype uses PM4Py as the core process mining library and combines it with pandas, NumPy, Plotly, NetworkX, PyYAML, Graphviz and OpenPyXL for data processing, visualization, network analysis, configuration handling and Excel input.

To provide a consistent execution environment, the prototype was packaged as a Docker container. The Docker image is based on Python 3.11-slim, installs Graphviz and the Python packages specified in the requirements file, and runs `process_mining_tool_pm4py.py` as its entry point. The input event log and generated report are handled through separate `/data` and `/out` volumes, while the SLA thresholds are supplied through an external YAML configuration file. This setup allows the analysis to be rerun without installing PM4Py and its supporting libraries directly on the host computer.

PM4Py is particularly suitable for this purpose. Berti et al. (2023) describe it as a Python library that provides a comprehensive array of process mining tools and integrates with other Python libraries. It supports process discovery, directly-follows graphs, Petri nets, token-based replay, alignments, social and organizational analysis, resource profiling and several event log and process model formats.

The prototype follows the architecture presented in Figure 1.

Figure 1: Architecture of the Python-based forensic process mining prototype



Source: Authors’ illustration.

The program accepts CSV, Excel and XES inputs, maps the key event log fields and converts data into the PM4Py structure. It generates variants, DFG frequencies and performance measures, a Petri-net-based model, token replay fitness and alignment indicators. It also calculates waiting times, service times, workload, handover-of-work, temporal heatmaps, centrality measures, Benford indicators, z-score outliers and SLA violations defined in YAML or JSON files.

The final output is an interactive HTML report containing tables and visualizations. The analytical modules are summarized in Table 2.

Table 2: Analytical modules of the Python prototype and forensic purpose

Module	Technical function	Forensic purpose
Data import and mapping	CSV, Excel or XES input; mapping of case, activity, timestamp and resource columns	prepares event logs for analysis and reduces dependency on one data format
Variant analysis	identifies the most frequent process variants	highlights dominant and rare process paths
DFG frequency and performance	calculates directly-follows relations and mean times	detects frequent transitions and delayed process segments
Petri net discovery and conformance	discovers a model and applies token replay or alignments	identifies deviations between observed and modelled behaviour
Bottleneck and service-time analysis	calculates waiting and activity execution times	detects suspicious delays or unusually short control performance
Resource and handover analysis	analyses workload, handovers and centrality	identifies unusual actor involvement and potential SoD concerns
Temporal heatmaps	shows activity distribution by day and hour	supports detection of unusual timing patterns
Financial analytics	applies Benford indicators and z-score outlier detection	connects process deviations with financially relevant anomalies
SLA rules	applies externally defined thresholds	identifies breaches of expected process timing
HTML reporting	creates interactive tables and visualizations	documents red flags for further investigation

Source: Authors' analysis.

Generative AI was used at two stages of the proof of concept. ChatGPT was used to assist in generating the simulated procurement event log and to support selected code-drafting, debugging and refinement tasks during development of the Python prototype. The authors retained control over the forensic logic, research design and SLA rules, executed the program in the Docker environment and reviewed the resulting tables and visualizations. All methodological decisions and interpretations of the detected red flags remained the responsibility of the authors.

The source code, configuration files and simulated event log used in the proof of concept are available from the corresponding author upon reasonable request.

4. Proof-of-Concept Case Study: Procurement Event Log Analysis

The proof-of-concept case study is based on a simulated procurement process event log. Procurement was selected because it is exposed to multiple fraud risks, including unauthorized purchases, supplier favoritism, manipulation of approval procedures, invoice-related irregularities, conflicts of interest and weaknesses in segregation of duties. The objective of the case study is not to prove the existence of fraud, but to demonstrate how a Python-based forensic process mining tool can transform event log data into structured red flags.

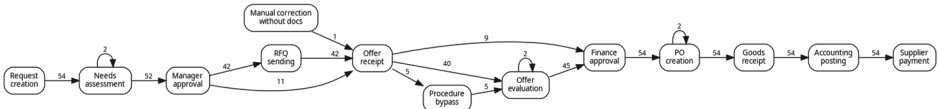
The event log contains 54 cases, 590 events and 14 activity types. Each event includes a case identifier, activity name, timestamps, employee, department, supplier, cost center, amount, value added tax, currency, procurement type, approval field, location and priority. Activities include regular procurement steps and exceptional activities such as procedure bypass, manual correction without documentation and repeated approval without reason.

The 54 procurement cases were generated synthetically with the assistance of ChatGPT and stored as an Excel event log containing the fields listed above. ChatGPT was used to create the case sequences, timestamps and case attributes needed for the proof-of-concept analysis. The generation process was not designed as a controlled fraud-detection experiment: no predefined problems, labelled anomalies or expected red-flag outcomes were deliberately inserted in advance. Exceptional activities appearing in the dataset were identified only after the prototype had processed the event log.

The SLA configuration defines a maximum case duration of 120 hours. It also defines maximum median waiting times for three critical transitions: 24 hours from finance approval to purchase order creation, 48 hours from goods receipt to accounting posting, and 72 hours from accounting posting to supplier payment. This configuration illustrates a key advantage of the Python approach: business rules can be specified externally and modified without changing the analytical code.

The generated report first provides basic metrics and process discovery results. The process model and directly-follows graph (DFG) visualization help the analyst understand the actual procurement flow. Variant analysis lists the most frequent sequences of activities and allows the analyst to identify rare paths. In a forensic context, rare variants do not automatically indicate fraud, but they represent useful starting points for further inquiry, especially when combined with other attributes such as supplier, amount, employee or control activity. Figure 2 presents a simplified DFG generated from the simulated event log.

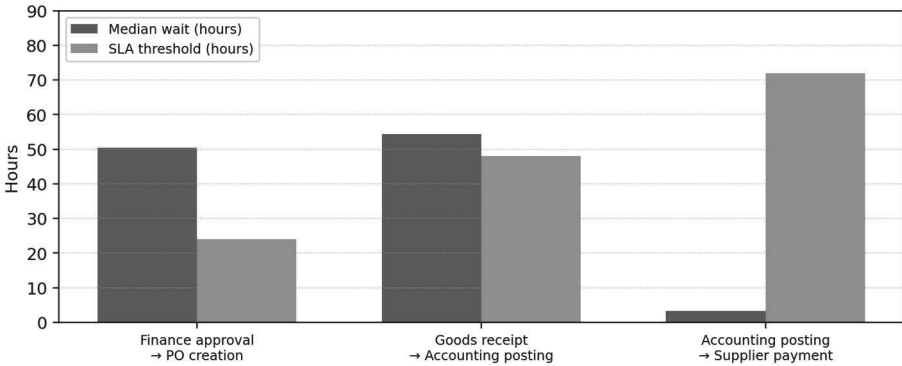
Figure 2: Simplified directly-follows graph of the simulated procurement process



Source: Authors' analysis based on the Python prototype.

The SLA module produced two types of findings. First, 29 cases exceeded the 120-hour maximum process duration. Second, two median waiting-time violations were identified: the transition from finance approval to purchase order creation had a median waiting time of 50.40 hours compared with a 24-hour threshold, and the transition from goods receipt to accounting posting had a median waiting time of 54.32 hours compared with a 48-hour threshold. The transition from accounting posting to supplier payment did not exceed the defined 72-hour threshold. These results do not demonstrate fraud, but they identify process segments and cases that require explanation. Figure 3 summarizes the SLA comparison.

Figure 3: SLA thresholds and median waiting times for selected procurement transitions



Source: Authors' analysis based on the Python prototype.

The resource module adds another forensic perspective. Workload, handover-of-work, centrality and temporal heatmaps may indicate whether one employee is disproportionately involved in critical activities, whether cases repeatedly move between the same actors, or whether activities occur at unusual times.

The financial module identified z-score outliers in amount and value added tax columns and calculated Benford indicators. These tests are not sufficient to infer manipulation, but they help combine process-level red flags with financially relevant deviations.

Validation was conducted ex post through a manual review of the analytical report generated by the prototype. The authors examined the reported tables and visualizations for internal consistency and plausibility in relation to the simulated

procurement process. Since the dataset contained no predetermined anomalies or ground-truth labels, this validation did not measure fraud-detection accuracy or false-positive and false-negative rates. Its purpose was to assess whether the prototype produced coherent and interpretable analytical results from the simulated event log.

The case study therefore demonstrates the main value of forensic process mining: it does not replace professional judgment, but it structures attention. It creates a map of cases, activities, actors and process segments that appear unusual according to predefined analytical logic. This enables forensic accountants and auditors to prioritize further procedures, request supporting documentation, interview process owners, test authorizations and examine related accounting records.

5. Discussion: Contribution, Limitations and Future Development

This paper makes four contributions. First, it positions forensic process mining as a specific application of process mining in forensic accounting. Existing literature has already shown that process mining is relevant for auditing, internal controls and fraud-related analytics (Föhr et al., 2025; Jans & Eulerich, 2022; Rodríguez-Quintero et al., 2021; Werner et al., 2021). This paper extends that discussion by focusing on red flags embedded in business event logs.

Second, the paper proposes a practical classification of forensic red flags in event logs. The classification includes control-flow, time-based, resource-based, case- or supplier-based, financial, conformance and behavioural indicators. This classification can help forensic accountants translate broad fraud risk concerns into concrete analytical tests.

Third, the paper demonstrates that Python and PM4Py can be used to develop a customized forensic process mining prototype. This is relevant because forensic accounting often requires domain-specific logic that may not be available in standard process mining dashboards. The prototype shows how process discovery, conformance checking, performance analysis, resource mining, SLA monitoring and financial analytics can be combined in one report.

Fourth, the paper briefly illustrates that generative AI can reduce technical barriers for forensic accountants by supporting code drafting and debugging. However, forensic logic, risk assessment and interpretation must remain the responsibility of qualified professionals.

The approach has several limitations. The case study relies on a single event log generated with the assistance of ChatGPT rather than on real organizational data; it therefore demonstrates feasibility rather than empirical accuracy in detecting fraud. No predefined anomalies or ground-truth fraud labels were embedded in the dataset, and the detected red flags were not benchmarked against known irregularities or findings identified by independent experts. Validation was limited to an ex post review of the internal consistency and plausibility of the analytical outputs. The

prototype is not a complete forensic software solution. Although the resource module can indicate unusual workload and handover patterns, it does not perform rule-based segregation-of-duties testing. Supplier-specific anomaly scoring and automated case-level risk scoring are also not implemented. The results depend on event log quality. If activities are missing, timestamps are inaccurate or relevant controls are performed outside the source system, the analysis may be incomplete. Werner et al. (2021) emphasize that process mining can inspect only transactions and controls recorded in source systems; activities outside those systems may be assessed only indirectly.

Privacy and confidentiality are additional limitations. Event logs may include personal data about employees, suppliers and customers; therefore, responsible use requires access control, data minimization, anonymization where appropriate, and clear governance over how red flags are interpreted.

Future development should extend the prototype into a more complete forensic exception reporting tool. Relevant extensions include rule-based segregation of duties testing, supplier-level variant analysis, authorization checks, case-level red flag scoring, machine learning anomaly detection, direct ERP integration and offline HTML reporting. Future research should validate the approach using real anonymized organizational data and compare detected red flags with documented audit or forensic findings.

Conclusion

Digital business systems generate event logs that can serve as important sources of forensic accounting insight. These logs record not only financial outcomes, but also the process through which business transactions were created, approved, modified, posted and paid. Forensic process mining extends the analytical scope of forensic accounting from transaction-level analysis to process-level digital evidence.

The paper conceptualized forensic process mining as the application of process mining techniques, forensic accounting logic and audit analytics to business event logs for the purpose of identifying red flags. A Python-based proof-of-concept prototype demonstrated how event logs can be analysed through process variants, directly-follows relations, conformance indicators, bottlenecks, service times, resource behaviour, SLA rules, temporal patterns and financial outlier tests. The procurement case study showed that such a tool can generate structured indicators for further audit or forensic procedures.

The main conclusion is that Python-based forensic process mining is feasible and practically relevant. It enables forensic accountants to customize analytical rules, combine process and financial attributes, and produce investigation-oriented reports. Generative AI may further lower the technical barrier for developing such tools, but it should be treated only as programming support. Detected red flags remain

indicators, not evidence of fraud. Professional judgment, documentation review and additional forensic procedures remain essential.

The findings must be interpreted within the limitations of the research design. The study relies on a single simulated procurement event log comprising 54 cases and 590 events generated with the assistance of ChatGPT, rather than on real organizational data. Because no predetermined anomalies or ground-truth labels were embedded, and validation was limited to an ex post review of the analytical outputs, the study does not establish fraud-detection accuracy. The current prototype also does not implement rule-based segregation-of-duties testing, supplier-specific anomaly scoring or automated case-level risk scoring. In addition, its findings depend on the completeness and accuracy of the event log. The results should therefore be interpreted as evidence of feasibility and as indicators for further investigation, rather than as validated fraud classifications.

Future research should validate this approach on real anonymized data, expand the prototype with segregation of duties and supplier-specific risk modules, and examine how forensic process mining can be embedded into internal audit, compliance and fraud risk management practices.

References

- Berti, A., van Zelst, S. J., & Schuster, D. (2023). PM4Py: A process mining library for Python. *Software Impacts*, 17, Article 100556. <https://doi.org/10.1016/j.simpa.2023.100556>
- da Silva, M. C., Tavares, G. M., Gritti, M. C., Ceravolo, P., & Barbon Junior, S. (2023). Using process mining to reduce fraud in digital onboarding. *FinTech*, 2(1), 120–137. <https://doi.org/10.3390/fintech2010009>
- Drakoulogkonas, P., & Apostolou, D. (2021). On the selection of process mining tools. *Electronics*, 10(4), Article 451. <https://doi.org/10.3390/electronics10040451>
- Eulerich, M., Huang, Q., Pawlowski, J., & Vasarhelyi, M. A. (2025). Using process mining as an assurance tool in the three-lines-model. *International Journal of Accounting Information Systems*, 56, Article 100731. <https://doi.org/10.1016/j.accinf.2025.100731>
- Föhr, T. L., Reichelt, V., Marten, K.-U., & Eulerich, M. (2025). A framework for the structured implementation of process mining for audit tasks. *International Journal of Accounting Information Systems*, 56, Article 100727. <https://doi.org/10.1016/j.accinf.2025.100727>
- Jans, M., & Eulerich, M. (2022). Process mining for financial auditing. In W. M. P. van der Aalst & J. Carmona (Eds.), *Process Mining Handbook* (pp. 445–467). Springer. https://doi.org/10.1007/978-3-031-08848-3_15

Khan, S., Parkinson, S., & Murphy, C. (2023). Context-based irregular activity detection in event logs for forensic investigations: An itemset mining approach. *Expert Systems with Applications*, 233, Article 120991. <https://doi.org/10.1016/j.eswa.2023.120991>

Rodríguez-Quintero, J. F., Sánchez-Díaz, A., Iriarte-Navarro, L., Maté, A., Marco-Such, M., & Trujillo, J. (2021). Fraud audit based on visual analysis: A process mining approach. *Applied Sciences*, 11(11), Article 4751. <https://doi.org/10.3390/app11114751>

van der Aalst, W. M. P. (2022). Process mining: A 360 degree overview. In W. M. P. van der Aalst & J. Carmona (Eds.), *Process Mining Handbook* (pp. 3–34). Springer. https://doi.org/10.1007/978-3-031-08848-3_1

Werner, M., Wiese, M., & Maas, A. (2021). Embedding process mining into financial statement audits. *International Journal of Accounting Information Systems*, 41, Article 100514. <https://doi.org/10.1016/j.accinf.2021.100514>

Wolfgang Raeder
Heilbronn University of Applied
Sciences, Faculty of Management
and Sales
ORCID: 0009-0006-5347-595X

Jelena Demko-Rihter
University of Novi Sad, Faculty of
Technical Sciences
ORCID: 0000-0001-7762-3095

DESIGN-LEVEL FORENSICS: USING INTELLIGENT AGENTS TO SPOT FRAUD VULNERABILITIES IN SAP CONFIGURATION

UDK 004.8:657.6
DOI <https://doi.org/10.54318/czf.2026.wr.012>

Abstract

Analyzing configurations extracted from enterprise IT systems such as SAP to identify design-level fraud risks and control weaknesses traditionally requires auditors and consultants with deep expertise in internal data structures and business logic. This dependency creates a bottleneck, particularly for less experienced practitioners. This practice-based design paper presents an approach, developed in cooperation with KPMG, that uses large language models (LLMs) to support and partially automate this forensic design analysis, thereby lowering the expertise threshold required for SAP configuration audits.

The solution follows a two-stage architecture. First, raw configuration data extracted from SAP systems is processed by topic-specific evaluation routines that structure and pre-aggregate relevant information. Second, the compact results are combined with tailored prompts and transmitted via a REST interface to a language model, which identifies potential fraud vulnerabilities and configuration weaknesses and aggregates its granular responses into control-specific, audit-ready findings. An exploratory pilot in three real audits indicated time savings of about 20–30%, driven mainly by reduced documentation effort.

Extending this architecture toward Agentic AI, the solution is embedded as a Model Context Protocol (MCP) server within an autonomous agent, which prepares the full workflow from data collection and analysis to the documentation of results, while approval and client communication remain with the auditor. The paper discusses the design, implementation, preliminary evaluation and professional safeguards of this approach and reflects on its implications for accessibility and scalability of design-level forensics in SAP environments.

Key words: SAP; Fraud Detection; Large Language Models; Agentic AI; IT Audit; Model Context Protocol; Forensic Data Analysis

JEL Classification: M42, M15, O33, K42, C88

Introduction

A well-designed internal control system sits at the center of fraud prevention, detection, and deterrence (Committee of Sponsoring Organizations of the Treadway Commission [COSO], 2013), and effective anti-fraud controls correlate with materially lower losses (Association of Certified Fraud Examiners [ACFE], 2026). Yet in SAP-based enterprise resource planning (ERP) environments, controls are only as good as the configuration beneath them: authorization roles, parameter settings and validation checks must be correctly assembled before the system's control capabilities become effective (Chuprunov, 2013a, 2013b). Given how vast and interdependent SAP's configurable settings are, judging whether a configuration amounts to a design-level fraud vulnerability is technically demanding work, and many control weaknesses trace back to an expertise gap among the staff configuring and reviewing these systems (Donelson, Ege, & McInnis, 2017).

Existing approaches to SAP fraud-risk assessment — manual review against control frameworks, or continuous auditing and monitoring tools — automate mainly the extraction and rule-based flagging of configuration data (Little & Best, 2003; Singh & Best, 2015). Interpreting whether a flagged setting constitutes a genuine fraud vulnerability still demands deep SAP expertise (Kuhn & Sutton, 2010), so assessment quality tracks reviewer experience. At the same time, large language models (LLMs) are recognized as a viable auditing and forensic-accounting technology (Vasarhelyi, Moffitt, Stewart, & Sunderland, 2023), and agentic architectures relying on standardized protocols such as the Model Context Protocol (MCP) are emerging (Ehtesham, Singh, Gupta, & Kumar, 2025; Schreyer, Gu, Moffitt, & Vasarhelyi, 2024). To our knowledge, no published approach combines LLM-based interpretation of SAP configuration data with an agentic, MCP-based architecture for design-level fraud vulnerability detection.

This paper is a practice-based design contribution with an exploratory evaluation of a software artefact developed in cooperation with KPMG. It addresses three design questions: (DQ1) How can an LLM-supported architecture reduce the expertise bottleneck in design-level fraud analysis of SAP configurations? (DQ2) How should the analytical work be divided between rule-based evaluation and LLM-based interpretation? (DQ3) Which preliminary benefits and limitations can be observed in practice?

Section 1 reviews related work. Section 2 presents the two-stage architecture and its agentic extension (DQ1, DQ2), and Section 3 its implementation. Section 4 reports the exploratory pilot (DQ3), Section 5 discusses safeguards and limitations, and the conclusion answers the design questions.

1. Background and Related Work

1.1 Fraud risk prevention in ERP/SAP environments

COSO's Internal Control – Integrated Framework comprises five components: control environment, risk assessment, control activities, information and communication, and monitoring (COSO, 2013). Within IT-supported environments, control activities split into IT general controls (ITGCs), covering access, change, and operations, and IT application controls (ITACs) embedded within business processes. In SAP systems, both rest on the authorization concept and its configuration settings, but the technology only supplies the capability for control. The effectiveness depends on correct configuration during customization (Chuprunov, 2013a, 2013b). Where configuration is incomplete or inconsistent with an organization's risk profile, control gaps emerge, and such weaknesses substantially raise the likelihood of undetected financial reporting fraud (Donelson et al., 2017).

1.2 Traditional audit approaches to SAP configuration analysis

Auditing SAP systems for fraud risk means reviewing configuration against established frameworks, either manually or with continuous auditing/monitoring (CA/CM) tools; both presuppose SAP-specific expertise to translate technical parameters into business-process risk (Kuhn & Sutton, 2010). Within ITGCs, audit programs typically cover access provisioning, change and transport management, and logging. Within ITACs, attention shifts to authorization structures such as segregation of duties (SoD). Little and Best (2003) propose a role-based framework for evaluating SoD within SAP R/3, showing that adequate SoD design has to be tailored to an organization's specific roles and transactions. Singh and Best (2015) show that continuous monitoring embedded in SAP ERP raises audit coverage, while its configuration and interpretation still hinge on experienced auditors. Traditional approaches thus provide well-established frameworks, but require SAP-experienced personnel for interpretation, precisely where control weaknesses tend to originate (Section 1.1).

1.3 Large language models in auditing and forensic accounting

Since late 2022, a growing body of research has examined LLMs' applicability to accounting, auditing, and fraud-detection tasks. Vasarhelyi et al. (2023) map the field, flagging open questions around reliability and professional judgment. Eulerich and Wood (2025) show how generative AI can support internal-audit planning, testing, and reporting, and Berger et al. (2023) present LLM-based automated regulatory compliance verification. Kirkos, Boskou, Chatzipetrou, Tiakas, and Spathis (2024) report mixed but improving performance of LLM-based financial statement fraud classification, while Jezierski, Kaltenpoth, Müller, and Weißenberger (2025) offer causal field evidence that LLM assistance improves audit quality. Taken together, this literature signals growing recognition of LLMs as viable

tools for auditing and forensic accounting. It remains, however, largely confined to text-based financial analysis; structured configuration data such as SAP authorization settings is hardly addressed.

1.4 Agentic AI and the Model Context Protocol (MCP)

Agentic AI pushes the LLM paradigm toward systems that plan, invoke external tools, and execute multi-step workflows with limited human intervention (Xi et al., 2023). MCP, introduced by Anthropic (2024), standardizes how such agents connect to external tools and data sources through a client–server design for secure tool invocation and context exchange. It is emerging as a *de facto* standard and is recommended as the first step of a phased adoption roadmap for agent interoperability (Ehtesham et al., 2025; Singh, Ehtesham, Kumar, & Khoei, 2025). Within auditing, Schreyer et al. (2024) demonstrate autonomous, multi-step journal-entry testing, while Hsieh (2025) turns to the governance implications of autonomous agents. Existing agentic-auditing work, however, concentrates on financial-statement and journal-entry tasks; applying MCP-based agentic architectures to ERP configuration data remains unexplored.

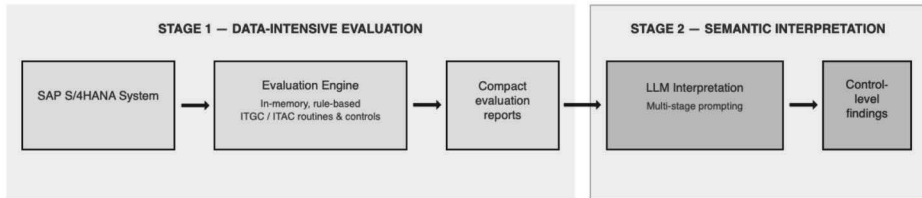
2. Solution Design

2.1 Overall architecture (two-stage approach)

SAP database tables are typically large in both record count and interrelation. External auditors or consultants usually cannot access the client's productive S/4HANA instance for data-protection reasons; data is therefore extracted and analyzed separately (Kogan, Alles, Vasarhelyi, & Wu, 2014), which also allows a standardized approach to be applied consistently across client systems, independent of company-specific system configurations and authorization concepts.

A central design decision concerns how LLMs are integrated into this evaluation. Passing SAP tables directly into an LLM prompt would exceed current context-window limits, and long contexts degrade well before the nominal limit (Liu et al., 2023). Letting the LLM generate and execute its own analysis code avoids this constraint, but is expected to be less efficient, since general-purpose code-execution environments rely on generic in-memory structures not optimized for large, relational SAP tables, and iterative code generation introduces additional latency. This solution therefore uses ISPICIO's standard analysis software, which provides fast in-memory processing tailored to SAP data and native LLM connectivity, to perform the evaluation routines (Stage 1). The LLM only interprets their compact output (Stage 2), as shown in Figure 1. This keeps prompts small and token-driven costs low, consistent with the trend of augmenting LLMs with external tools (Schick et al., 2023).

Figure 1: Two-stage architecture: Stage 1 reduces raw SAP data to compact results before Stage 2 is applied



Source: Authors' research

2.2 Stage 1: Topic-specific data preparation and evaluation routines

Relevant database tables are extracted from the SAP system (S/4HANA or ECC) either via SAP's Remote Function Call (RFC) interface or via a custom ABAP add-on, an established pattern in IS audit comparable to embedded audit modules used to obtain ERP data for independent analysis while limiting interference with production systems (Debreceeny, Gray, Ng, Lee, & Yau, 2005).

The data is then processed using predefined evaluation routines covering ITGC areas (access, change, and operations; ISACA, 2019; Institute of Internal Auditors [IIA], 2012) and ITAC aspects such as automated tolerance checks. Individually configurable rules for authorizations and parameters allow the evaluation logic to be tailored to a specific control environment. Rules are grouped into higher-level controls according to a template control set, and the resulting reports form the input to Stage 2.

2.3 Stage 2: LLM-based interpretation

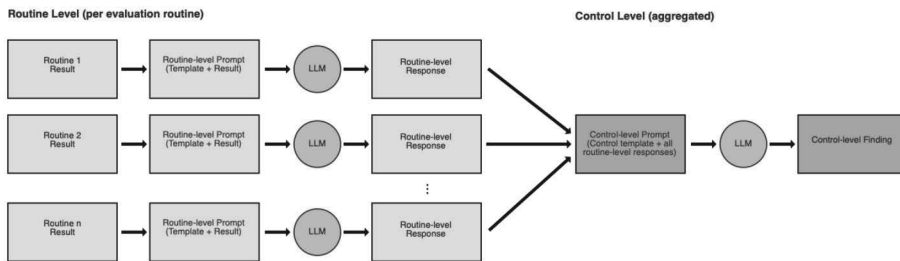
Interpreting and documenting findings from Stage 1's reports has traditionally required detailed SAP experience. Here, this task is delegated to the LLM, reflecting a broader shift toward AI-supported interpretation that changes, without eliminating, the expertise profile required of human reviewers (Kokina & Davenport, 2017).

The LLM is connected to the ISPICIO analysis software via a Representational State Transfer (REST) interface using the LangChain4j library, so that different LLMs like ChatGPT, Gemini or Claude, as well as local LLMs, can be used interchangeably depending on cost, latency, or compliance requirements.

For each evaluation routine and each control, a dedicated prompt template defines the information the LLM needs for correct interpretation: the underlying rule's meaning, relevant thresholds, and expected output structure. Templates are managed within ISPICIO alongside the routines and controls themselves, keeping interpretive and analytical logic jointly versioned.

Interpretation proceeds hierarchically (Figure 2). First, a separate prompt combining the routine's template with its specific result is sent to the LLM per evaluation routine. Second, the routine-level responses of a control are combined with the control-level template into an aggregated prompt, whose response yields the control-level finding. For instance, several authentication rules (e.g., password complexity parameters) are first interpreted separately before being synthesized into one control-level assessment for authentication to the SAP system. This prompt chaining (Wu, Terry, & Cai, 2022) keeps prompts short, reducing the risk of hallucinated or inconsistent output (Ji et al., 2023), and keeps each intermediate response individually inspectable, supporting explainability (Arrieta et al., 2020).

Figure 2: Multi-stage prompt generation and aggregation



Source: Authors' research

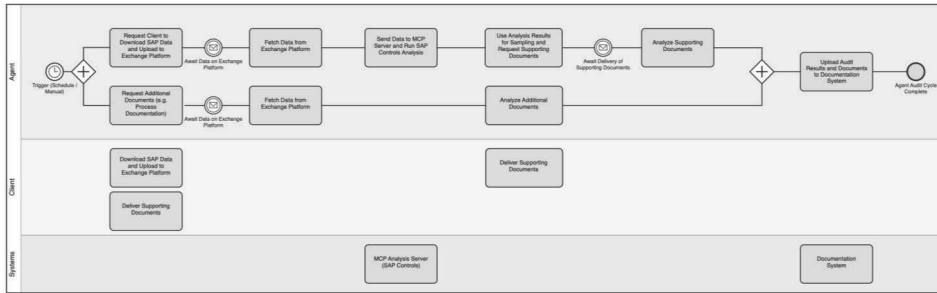
2.4 Extension toward Agentic AI: embedding the solution as an MCP server

The process in Sections 2.1–2.3 still requires a human user to coordinate each step. In the agentic extension, this coordination is delegated to a software agent that plans, invokes tools, observes results, and adapts subsequent actions, using an MCP server as the interface to the ISPICIO software. MCP standardizes how AI applications connect to external tools and data sources in a provider-agnostic way rather than through bespoke, one-off integrations (Anthropic, 2024), so that the agent can invoke the evaluation routines as part of a larger plan.

The resulting workflow comprises three stages (Figure 3). First, the agent prepares the data request to the relevant client contact, monitors a data-exchange platform, downloads the data via browser automation once available, and checks it for completeness and correctness. Second, the agent invokes the MCP server, which provides functions to upload the SAP data, the control and rule set, and the report templates, to start the analysis with selected rules and a defined review period, and to download the result files (HTML, CSV, and Markdown). The agent processes these files further, for instance by defining a random sample and preparing requests for supporting documents. Third, the agent prepares the documentation of procedures performed and the findings for client communication. The agent prepares all steps, whereas the release of results and the sending of e-mails remain with the

auditor, consistent with the emphasis of the auditing literature on continued human review and the ability to override an agent's actions and conclusions (Kokina & Davenport, 2017).

Figure 3: Agent-driven workflow (BPMN): the agent coordinates data collection, MCP-based analysis, sampling, and documentation



Source: Authors' research

3. Implementation

While the architecture in Section 2 is largely implementation-agnostic, a brief technical description is given here for completeness.

The application is implemented in Java. The MCP server exposing the ISPICIO routines (Section 2.4) is built using Spring Boot, a widely adopted framework for modular Java backend systems that aligns naturally with MCP's tool-based invocation model. The server is deployed as a Docker container, encapsulating the application and its dependencies into a portable unit that behaves consistently across host environments. The AI agent integrating the MCP server is currently implemented with Visual Studio Code and covers all steps shown in Figure 3, from data collection to MCP-based analysis, sampling, and documentation.

4. Exploratory Pilot

The non-agentic architecture (Sections 2.1–2.3) was piloted with KPMG in three real audit engagements, applying an ITGC rule set (access, change, and operations) and GPT-4o. To avoid additional costs, the task was not performed once manually and once with LLM support. Instead an experienced auditor validated the LLM responses and assessed the resulting time savings based on professional experience. The estimate includes the time for reviewing and selecting the LLM output, but not the final review of the audit in the documentation system, which takes place subsequently. The figure of 20–30% rests on a limited number of engagements and should not be read as a precise, generalizable estimate. Table 1 summarizes the evaluation dimensions, observations, and limitations.

Table 1: Exploratory pilot: evaluation dimensions, observations, and limitations

Dimension	Observation	Limitation / basis
Scope	Three real audit engagements with KPMG; ITGC rule set; GPT-4o	Small, non-representative sample
Comparison	No parallel manual execution; time savings estimated by an experienced auditor	No measured baseline; estimate based on professional experience
Time savings	Approx. 20–30% (estimated)	Estimate includes review and selection of LLM output; excludes the subsequent final audit review
Accuracy	Comparable to manual work, validated by an experienced auditor	Preliminary practitioner observation; no formal accuracy measure
Repeatability	Good over short periods; declining over several months	Provider-side changes in model behavior; local LLMs as possible remedy
Supporting documents	Scanned PDFs require OCR	Additional effort and risk of recognition errors

Source: Authors' research

In addition, supporting documents obtained for sample testing (Section 2.4) must contain machine-readable text; scanned PDFs require an additional optical character recognition (OCR) step, which adds effort and, depending on scan quality, a risk of recognition errors that can offset part of the efficiency gains.

Overall, the main benefit lies in reduced documentation effort, since the LLM provides pre-formulated text at routine and control level that auditors can adopt or refine rather than draft from scratch — consistent with evidence that generative AI's productivity gains often stem from reducing drafting effort (Noy & Zhang, 2023). Tests of the agentic MCP integration (Section 2.4) will start in autumn 2026.

5. Discussion

Faster and more consistent evaluation of ITGC and authorization data supports a shift toward more frequent control monitoring rather than purely periodic review (Vasarhelyi, Alles, & Williams, 2010), which is relevant for fraud risk, given that access-related weaknesses tie back to the opportunity component of fraud. The standardized, tool-supported approach also permits a uniform methodology across

client environments. Lowering the SAP-specific expertise threshold broadens access to such reviews and shortens ramp-up time for less experienced staff.

This lowered threshold, though, carries risks: automation bias, i.e., acceptance of automated output without sufficient scrutiny (Parasuraman & Manzey, 2010), potentially amplified where reviewers lack the domain expertise to assess LLM findings critically (Commerford, Dennis, Joe, & Wang, 2022), and a long-term erosion of independent expertise through reduced hands-on engagement (Arnold & Sutton, 1998). Lowering the threshold should therefore enable scale and accessibility, not substitute for maintaining sufficient in-house expertise. Moreover, the multi-stage prompt design mitigates, without eliminating, hallucination risk (Ji et al., 2023), and reliance on externally hosted models introduces provider dependency affecting reproducibility.

These risks are addressed by professional safeguards. LLM output is always treated as a proposal; the decision on which findings are used remains with the auditor. All intermediate routine-level responses are retained, so that each control-level finding remains traceable to its inputs, and rules, controls, and prompt templates are versioned jointly (Section 2.3). The provider-agnostic integration allows processing to be restricted to local models where client confidentiality requires it. In the agentic workflow, the agent can only invoke the predefined MCP functions, and both the release of findings into the audit documentation and any e-mail to the client require the auditor's approval. The system thus augments rather than replaces professional judgment.

Key limitations remain. The pilot is small and lacks a systematic, independent assessment of accuracy across a large number of cases. The architecture covers only the SAP application layer, not the operating system or database layer that established ITGC frameworks treat as jointly relevant (IIA, 2012; ISACA, 2019), and the quality of findings is bounded by the human-curated rule base and prompt templates, which the LLM does not correct.

Conclusion and Future Work

This paper presented a two-stage architecture for LLM-supported analysis of SAP configuration data and its extension into an agentic, MCP-integrated workflow. Regarding DQ1, separating rule-based evaluation from LLM-based interpretation lowers the SAP-specific expertise required to interpret and document configuration findings, and the agentic MCP extension further reduces coordination effort. Regarding DQ2, data-intensive computation belongs to the specialized analytical engine, whereas the LLM interprets compact results through short, hierarchical prompts that remain traceable. Regarding DQ3, the pilot in three real audits indicated time savings of about 20–30%, driven mainly by reduced documentation effort, with accuracy comparable to manual work, but declining repeatability over time and additional effort for documents without machine-readable text.

Future work will evaluate the agentic workflow and assess accuracy systematically across a larger number of engagements. Further, the agent could retrieve an organization's procedural documentation, such as its operations and authorization concepts, to compare the actual SAP configuration against the documented control design, and the evaluation routines and prompt templates could be extended to operating-system and database-level data and to other ERP systems.

References

- Anthropic. (2024). Introducing the Model Context Protocol. <https://www.anthropic.com/news/model-context-protocol>
- Arnold, V., & Sutton, S. G. (1998). The theory of technology dominance: Understanding the impact of intelligent decision aids on decision maker's judgments. *Advances in Accounting Behavioral Research*, 1, 175–194.
- Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Benetot, A., Tabik, S., Barbado, A., ... Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115. <https://doi.org/10.1016/j.inffus.2019.12.012>
- Association of Certified Fraud Examiners. (2026). Occupational fraud 2026: A report to the nations. <https://www.acfe.com/-/media/files/acfe/pdfs/rtnn/2026/2026-report-to-the-nations.pdf>
- Berger, A., Hillebrand, L., Leonhard, D., Deußner, T., De Oliveira, T. B. F., Dilmaghani, T., Khaled, M., Kliem, B., Loitz, R., Bauckhage, C., & Sifa, R. (2023). Towards automated regulatory compliance verification in financial auditing with large language models. In 2023 IEEE International Conference on Big Data (BigData) (pp. 4626–4635). IEEE. <https://doi.org/10.1109/BigData59044.2023.10386518>
- Chuprunov, M. (2013a). IT general controls in SAP ERP. In Auditing and GRC automation in SAP (pp. 131–163). Springer. https://doi.org/10.1007/978-3-642-35302-4_6
- Chuprunov, M. (2013b). General application controls in SAP ERP. In Auditing and GRC automation in SAP (pp. 165–187). Springer. https://doi.org/10.1007/978-3-642-35302-4_7
- Commerford, B. P., Dennis, S. A., Joe, J. R., & Wang, J. (2022). Complex estimates and auditor reliance on artificial intelligence. *Journal of Accounting Research*, 60(2), 537–575. <https://doi.org/10.1111/1475-679X.12407>
- Committee of Sponsoring Organizations of the Treadway Commission. (2013). Internal control – Integrated framework. COSO.
- Debreceeny, R., Gray, G. L., Ng, J. J.-J., Lee, K. S. P., & Yau, W. F. (2005). Embedded audit modules in enterprise resource planning systems: Implementation

and functionality. *Journal of Information Systems*, 19(2), 7–27. <https://doi.org/10.2308/jis.2005.19.2.7>

Donelson, D. C., Ege, M. S., & McInnis, J. M. (2017). Internal control weaknesses and financial reporting fraud. *Auditing: A Journal of Practice & Theory*, 36(3), 45–69. <https://doi.org/10.2308/ajpt-51608>

Ehtesham, A., Singh, A., Gupta, G. K., & Kumar, S. (2025). A survey of agent interoperability protocols: Model Context Protocol (MCP), Agent Communication Protocol (ACP), Agent-to-Agent Protocol (A2A), and Agent Network Protocol (ANP) (arXiv:2505.02279). arXiv. <https://doi.org/10.48550/arXiv.2505.02279>

Eulerich, M., & Wood, D. A. (2025). A demonstration of how ChatGPT and generative AI can be used in the internal auditing process. *Journal of Emerging Technologies in Accounting*, 22(2), 47–77. <https://doi.org/10.2308/jeta-2023-041>

Hsieh, S.-F. (2025). Agentic AI and human-AI collaboration in auditing: Reframing roles, accountability, and governance. *International Journal of Computer Auditing*, 7(1), 57. <https://doi.org/10.53106/256299802025120701006>

Institute of Internal Auditors (IIA). (2012). GTAG 1: Information technology risk and controls (2nd ed.). Global Technology Audit Guide, IIA.

ISACA. (2019). COBIT 2019 framework: Governance and management objectives. ISACA.

Jezierski, M., Kaltenpoth, S., Müller, O., & Weißenberger, B. E. (2025). The effect of large language models on audit quality: Causal evidence from a randomized field experiment. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.6095166>

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., ... Fung, P. (2023). Survey of hallucination in natural language generation. *ACM Computing Surveys*, 55(12), Article 248. <https://doi.org/10.1145/3571730>

Kirkos, E., Boskou, G., Chatzipetrou, E., Tiakas, E., & Spathis, C. (2024). Exploring the boundaries of financial statement fraud detection with large language models. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.4842962>

Kogan, A., Alles, M. G., Vasarhelyi, M. A., & Wu, J. (2014). Design and evaluation of a continuous data level auditing system. *Auditing: A Journal of Practice & Theory*, 33(4), 221–245. <https://doi.org/10.2308/ajpt-50844>

Kokina, J., & Davenport, T. H. (2017). The emergence of artificial intelligence: How automation is changing auditing. *Journal of Emerging Technologies in Accounting*, 14(1), 115–122. <https://doi.org/10.2308/jeta-51730>

Kuhn, J. R., Jr., & Sutton, S. G. (2010). Continuous auditing in ERP system environments: The current state and future directions. *Journal of Information Systems*, 24(1), 91–112. <https://doi.org/10.2308/jis.2010.24.1.91>

- Little, A., & Best, P. J. (2003). A framework for separation of duties in an SAP R/3 environment. *Managerial Auditing Journal*, 18(5), 419–430. <https://doi.org/10.1108/02686900310476882>
- Liu, N. F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., & Liang, P. (2023). Lost in the middle: How language models use long contexts. <https://arxiv.org/abs/2307.03172>
- Noy, S., & Zhang, W. (2023). Experimental evidence on the productivity effects of generative artificial intelligence. *Science*, 381(6654), 187–192. <https://doi.org/10.1126/science.adh2586>
- Parasuraman, R., & Manzey, D. H. (2010). Complacency and bias in human use of automation: An attentional integration. *Human Factors*, 52(3), 381–410. <https://doi.org/10.1177/0018720810376055>
- Schick, T., Dwivedi-Yu, J., Dessi, R., Raileanu, R., Lomeli, M., Zettlemoyer, L., ... Scialom, T. (2023). Toolformer: Language models can teach themselves to use tools. [arXiv:2302.04761](https://arxiv.org/abs/2302.04761).
- Schreyer, M., Gu, H., Moffitt, K., & Vasarhelyi, M. A. (2024). Artificial intelligence agentic auditing [Working paper]. SSRN. <https://doi.org/10.2139/ssrn.4909147>
- Singh, A., Ehtesham, A., Kumar, S., & Khoei, T. T. (2025). A survey of the model context protocol (MCP): Standardizing context to enhance large language models (LLMs) [Preprint]. Preprints.org. <https://doi.org/10.20944/preprints202504.0245.v1>
- Singh, K., & Best, P. J. (2015). Design and implementation of continuous monitoring and auditing in SAP enterprise resource planning. *International Journal of Auditing*, 19(3), 307–319. <https://doi.org/10.1111/ijau.12051>
- Vasarhelyi, M. A., Alles, M. G., & Williams, K. T. (2010). Continuous assurance for the now economy. Institute of Chartered Accountants in Australia.
- Vasarhelyi, M. V., Moffitt, K. C., Stewart, T., & Sunderland, D. (2023). Large language models: An emerging technology in accounting. *Journal of Emerging Technologies in Accounting*, 20(2), 1-10. <https://doi.org/10.2308/jeta-2023-047>
- Wu, T., Terry, M., & Cai, C. J. (2022). AI Chains: Transparent and controllable human-AI interaction by chaining large language model prompts. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems (CHI '22) (Article 385, pp. 1–22). ACM. <https://doi.org/10.1145/3491102.3517582>
- Xi, Z., Chen, W., Guo, X., He, W., Ding, Y., Hong, B., ... Gui, T. (2023). The rise and potential of large language model based agents: A survey. [arXiv:2309.07864](https://arxiv.org/abs/2309.07864).

Jovana Stojanović
University of Niš – Innovation Center
ORCID: 0000-0003-0190-3802

Dejan Spasić
University of Belgrade –
Faculty of Economics and Business
University of Niš –
Faculty of Economics
ORCID: 0000-0002-8540-579X

INTERCONNECTION BETWEEN FINANCIAL DISTRESS AND THE RISK OF FRAUDULENT FINANCIAL REPORTING: EVIDENCE FROM BANKRUPT COMPANIES IN SERBIA

UDK 657.6:343.53

DOI <https://doi.org/10.54318/czf.2026.js.013>

Abstract

Turbulent market conditions, together with the consequences of systemic events that have weakened the Serbian economy, indicate the need for continuous monitoring of companies' operations. Since the consequences of crises are most often reflected in financial performance, the literature places particular emphasis on bankruptcy prediction models based on accounting indicators. However, the reliability of such models may be limited when financial statements, owing to the complexity of business operations and the ingenuity of fraud perpetrators, fail to faithfully represent companies' financial position and earning capacity. Therefore, to strengthen financial reporting transparency and confidence in the accounting profession, the application of bankruptcy prediction models should be accompanied by an assessment of financial reporting quality. This paper examines the applicability of the Altman Z"-score model and the five-variable Beneish M-score model using a sample of 40 Serbian companies for which bankruptcy proceedings were initiated between 2017 and 2024. The aim is to assess financial distress and the risk of fraudulent financial reporting prior to bankruptcy proceedings, with particular attention to the complementary information provided by the two models. The results indicate an increase in the prevalence of financial distress as bankruptcy approached, with the Altman model classifying 67.50% of companies into the distress zone two years before bankruptcy and 85.00% one year before it. The Beneish model, in turn, classified 17.50% of companies as potential manipulators, with warning signals associated with disproportionate movements in receivables relative to revenues. Taken together, the results show that most companies in the distress zone did not exhibit a Beneish manipulation signal, and no statistically significant association was identified between the two models' classifications. The findings may be useful to investors, financial institutions, regulators, and agencies involved in bankruptcy proceedings and fraud prevention, as well as researchers developing models tailored to specific national contexts.

Key words: bankruptcy prediction, forensic accounting, financial statements, fraud indicators, Z"-score, M-score

JEL Classification: G33, K13, M41

Introduction

Although business crisis can manifest itself in various ways, accounting data in publicly available financial statements stand for the most important source of information for confirming a company's financial distress. However, subjectivity in the assessment of individual items, the choice of accounting policies and manipulative practices may result in reported performance and financial position failing to faithfully reflect the underlying economic reality, particularly under conditions of crisis. When a company faces financial distress, its financial statements should allow for the timely identification of deteriorating solvency and an increased risk of bankruptcy. At the same time, however, management may face increased incentives to report more favorable results in order to meet investors' expectations, facilitate access to additional financing and preserve the company's legitimacy. All practices through which management misleads stakeholders regarding a company's financial performance and economic health by creating the appearance that its earnings, cash flows and financial position are more favorable than they actually are referred to as financial manipulation (Schilit, Perler & Engelhart, 2018). Subsequent disclosure of the company's true financial condition may result in a dramatic decline in its value and ultimately contribute to its bankruptcy. The Deloitte Forensic Center (2008) points to the link between fraudulent financial reporting and bankruptcy as well, claiming that companies facing bankruptcy are more likely to engage in financial statement manipulation. Therefore, these two risks should be assessed simultaneously through the application of appropriate models.

The application of financial distress prediction models is particularly important for Serbian companies, which are continuously exposed to liquidity, solvency and bankruptcy risks (Malinić, 2016; Malinić, 2023). The fact that, on average, approximately 5% of companies are subject to bankruptcy proceedings each year points to the seriousness of the problem. Although, according to the Bankruptcy Supervision Agency, the number of such proceedings in Serbia has not increased dramatically compared to developed countries (Muñoz-Izquierdo, Camacho-Miñano & Pascual-Ezama, 2017), the resulting losses are still measured in millions and have serious consequences for companies, stakeholders and society as a whole.

For the reasons mentioned above, the aim of the study is to assess financial distress and the risk of fraudulent financial reporting prior to bankruptcy proceedings, with particular attention to the complementary information provided by the Altman Z"-score model and the five-variable Beneish M-score model. In line with this aim, the following research questions were formulated:

Q1. To what extent did the Altman Z"-score model indicate the presence of financial distress in the years preceding the initiation of bankruptcy proceedings?

Q2. What proportion of the sampled companies was classified as potential manipulators by the Beneish M-score model prior to bankruptcy proceedings?

Q3. What differences in the individual Beneish indicators were observed between companies classified as potential manipulators and potential non-manipulators?

Q4. To what extent did financial distress signals from the Altman Z"-score model coexist with potential manipulation signals from the Beneish M-score model, and was there a statistically significant association between the two classifications?

To answer these questions, we use a sample of 40 companies against which bankruptcy proceedings were initiated in Serbia between 2017 and 2024.

The Altman and Beneish models have previously been applied separately in studies conducted in Serbia (Muminović, Pavlović & Cvijanović, 2012; Dimitrijević, Milutinović & Stanković, 2024; Malinić & Ribić, 2025) and other countries (Carton & Hofer, 2006; Chan & Landry, 2019). Studies in which the two models were applied simultaneously have generally focused on individual companies that went bankrupt and with confirmed fraudulent financial reporting (Mahama, 2015; MacCarthy, 2017). Even when the samples included a larger number of companies, they were not structured according to the bankruptcy status (Özari, Can & Demirkale, 2025; Svoboda, Berková, Pražáková & Vejsadová Dryjová, 2025). To the best of the authors' knowledge, the Altman and Beneish models have not been applied simultaneously to a group of companies undergoing bankruptcy proceedings. The contribution of this study therefore lies in addressing this gap while considering the complementary informational value of the two models.

To achieve the research objective, the article is structured as follows. Following the Introduction, the article reviews the financial reporting environment in Serbia and the relevant theoretical background. It then describes the methodology and data before presenting and discussing the results. The limitations of the study are addressed next, and the article concludes by summarizing the key findings and suggesting directions for future research.

1. Research context

1.1. The financial reporting environment in Serbia

Analyses of the performance of the Serbian economy in the periods 2007-2014 and 2017-2021 indicate that domestic companies have been continuously exposed to liquidity and solvency risks and, consequently, to the risk of bankruptcy (Malinić, 2016; Malinić, 2023). This situation is fundamentally driven by inadequate profitability and weak competitiveness, which constrain investment activity and hinder economic growth. At the same time, low operating efficiency, underutilization of production capacities, increasing indebtedness and high financial expenses further deteriorate companies' financial health. This is compounded by a low level of financial discipline, whereby the failure of some companies to meet their obligations can have a cascading effect on financially more stable companies. The scale of the problem is also reflected in the fact that approximately 20% of companies are over-indebted each year. Nevertheless, 5,394 companies were subject to

bankruptcy proceedings in 2024, accounting for just 20% of all over-indebted companies. This discrepancy can partly be explained by the duration and costs of bankruptcy proceedings, which create incentives for the out-of-court restructuring of debtor–creditor relationships. “Companies that are not viable but survive due to indiscriminate government assistance, thereby reducing the lending capacity available to healthy companies” pose an additional risk (Malinić, 2023, p. 6).

Business difficulties may often be closely associated with the quality, i.e. the reliability of financial statements. Formal verification of deviations from accounting standards comes from auditors’ reports when entities are subject to statutory audit requirements. The Vučković Milutinović (2019) study shows that, during the 2015–2017 period, 30.4% of companies listed on the Belgrade Stock Exchange received modified audit opinions, most commonly due to the overstatement of assets and understatement of liabilities, thereby distorting the picture of their financial position and earning capacity. Adverse opinions and the recurrence of modified opinions point to insufficient management incentives to remedy identified deficiencies. This situation may be associated with the underdeveloped capital market, poor demand for high-quality information, the marginalization of the accounting and auditing profession, regulatory instability and the lack of rigorous oversight of regulatory compliance (Dimitrijević, Obradović & Milutinović, 2018). Although such deficiencies do not necessarily constitute fraud, they represent warning signals that should not be disregarded.

1.2. Theoretical background – A literature review

Although various forms of financial difficulty may indicate a departure from the condition of a financially healthy company, risks associated with financial distress and bankruptcy are considered to be among the most severe. Early attempts to classify companies according to bankruptcy risk were based on univariate models, while multivariate approaches were subsequently developed. These include the Altman and Taffler models, based on multivariate discriminant analysis, as well as the Ohlson and Zmijewski models, based on logit and probit analysis, respectively. Although traditional bankruptcy prediction models are still used (Soukal et al., 2024), their applicability is constrained by the stringent statistical assumptions underlying them. Contemporary models, such as neural networks and decision trees, offer the possibility of overcoming some of the limitations of traditional approaches.

Crisis and bankruptcy are not exclusively the result of objective factors, such as technological and structural changes, monetary trends, global economic disruptions, recessions or declining activity in particular industries; they may also be the result of deliberate actions undertaken by owners and/or management to the detriment of creditors and other stakeholders (Škarić-Jovanović & Spasić, 2022). Such actions may be motivated by the pursuit of personal gain and the desire to maintain the appearance of rapid corporate growth, but also by attempts to conceal an already deteriorating financial condition. The predictive ability of models based on financial information has weakened over time (Singh & Mishra, 2016), which may, among

other factors, be attributed to the expanding scope for managerial discretion in financial reporting. Accounting regulations permit different approaches to the recognition and measurement of individual financial statement items, thereby affecting reported profit and equity. This flexibility enables earnings management, which is not inherently unethical but may develop into fraudulent financial reporting when used to intentionally misrepresent or omit financial information (Velte, 2023). Although fraudulent financial reporting accounted for only 6% of the 2,402 cases of fraud analyzed, it resulted in the largest median loss, amounting to USD 1 million per case (Association of Certified Fraud Examiners, 2026), which points to the seriousness of the problem.

In an effort to combat fraud, various approaches have been developed, ranging from conventional statistical models to machine learning and data mining techniques. Traditional models, such as the Beneish M-score, Dechow F-score and logistic regression, provide a useful basis for identifying manipulation; however, their application may be constrained by statistical assumptions and their predominant reliance on quantitative data. Contemporary approaches seek to overcome these limitations by integrating financial, non-financial and textual information, thereby enabling the identification of unusual and contradictory patterns that may indicate fraud (Compagnino et al., 2025). The application of artificial intelligence further contributes to the automated and continuous detection of suspicious patterns and to higher quality of financial reporting (Wang, Zhang & Han, 2025).

2. Methodology

The Altman model was selected for this study as one of the most widely used traditional models based on financial indicators. It summarizes a company's exposure to financial distress and bankruptcy risk in a single score that combines several dimensions of risk. Although the model was developed in the United States (US) in the second half of the twentieth century, studies have confirmed its applicability in other economic settings, including Serbia (Muminović et al., 2012; Muñoz-Izquierdo, Laitinen, Camacho-Miñano & Pascual-Ezama, 2020). Another reason for choosing the model was the availability of variants designed for samples with different characteristics. The original Z-score model was developed using data from publicly traded manufacturing companies, whereas the Z'-score model was adapted for private companies by replacing the market value of equity with its book value. Our study uses the Z''-score model, which was designed for non-manufacturing companies. This version excludes the sales-to-total-assets ratio to reduce the influence of industry-specific factors. Since it can be applied to both publicly traded and private companies, the model is well suited to the composition of the sample used in our study. The Altman Z''-score model is expressed as follows (Altman, Hartzell & Peck, 1995):

$$Z''\text{-score} = 6.56X_1 + 3.26X_2 + 6.72X_3 + 1.05X_4$$

where:

- X_1 – working capital / total assets;
 X_2 – retained earnings / total assets;
 X_3 – earnings before interest and taxes (EBIT) / total assets;
 X_4 – book value of equity / book value of total debt.

Once the Z"-score has been calculated, companies are classified into one of three risk zones. A Z"-score above 2.60 places a company in the safe zone, a score between 1.10 and 2.60 in the *grey zone*, and a score below 1.10 in the distress zone, indicating a high model-assessed risk of bankruptcy.

The Beneish model (Beneish, 1999) was selected for this study because it uses publicly available data, provides a defined classification threshold and supports the preliminary identification of potential financial reporting irregularities. The model has been shown to be useful for assessing the risk of fraudulent financial reporting (Dyck, Morse & Zingales, 2024) and identifying *red flags* across a range of countries (Özari et al., 2025; Svoboda et al., 2025), including Serbia (Spasić & Stojanović, 2026). It flagged an elevated risk of fraudulent reporting at Enron one year before the first public warnings by financial analysts (Beneish, Lee & Nichols, 2013).

The Beneish model produces a single score (M-score), which is compared against a cut-off value to identify potential warning signals of fraudulent financial reporting. The model is based on financial indicators (variables) that measure year-over-year changes in financial statement items. Both five-variable (Beneish M-5 model) and eight-variable (Beneish M-8 model) versions are used in the literature. This study applies the Beneish M-5 model, which is calculated as follows:

$$\text{M-score (5)} = -6.065 + 0.823 \times \text{DSRI} + 0.906 \times \text{GMI} + 0.593 \times \text{AQI} + 0.717 \times \text{SGI} + 0.107 \times \text{DEPI}$$

The calculation of the variables is presented in Table 1.

Table 1: Definition of study variables

Variables	Measurement
Days' sales in receivables index (DSRI)	$(\text{receivables}_t / \text{sales}_t) / (\text{receivables}_{t-1} / \text{sales}_{t-1})$
Gross margin index (GMI)	$[(\text{sales}_{t-1} - \text{cost of goods sold}_{t-1}) / \text{sales}_{t-1}] / [(\text{sales}_t - \text{cost of goods sold}_t) / \text{sales}_t]$
Asset quality index (AQI)	$[1 - (\text{current assets}_t + \text{net PP\&E}_t) / \text{total assets}_t] / [1 - (\text{current assets}_{t-1} + \text{net PP\&E}_{t-1}) / \text{total assets}_{t-1}]$
Sales growth index (SGI)	$\text{sales}_t / \text{sales}_{t-1}$
Depreciation index (DEPI)	$[\text{depreciation}_{t-1} / (\text{depreciation}_{t-1} + \text{net PP\&E}_{t-1})] / [\text{depreciation}_t / (\text{depreciation}_t + \text{net PP\&E}_t)]$

Source: Beneish (1999)

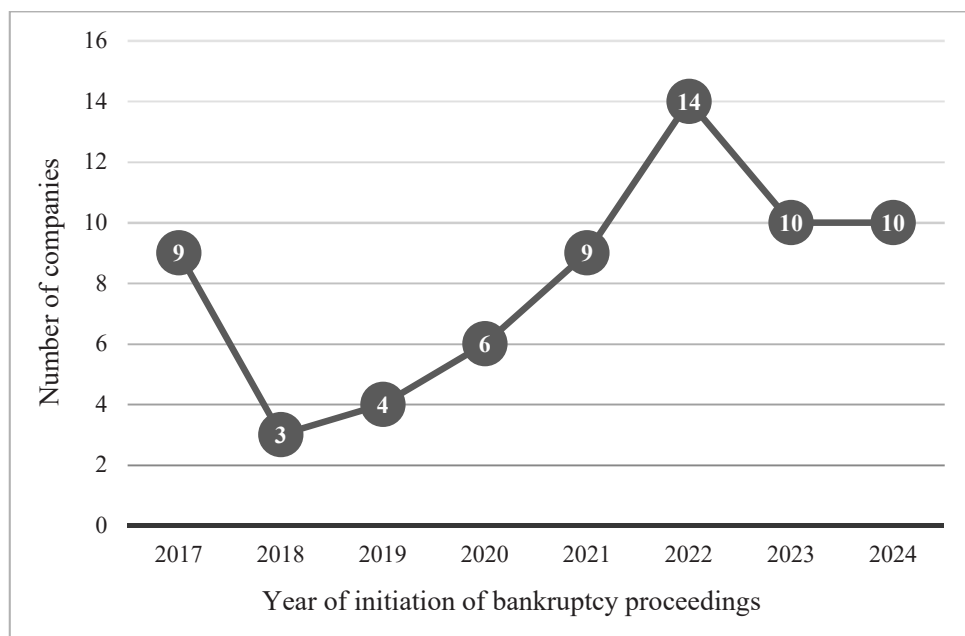
Although the M-score provides a basis for classifying companies, examining its individual variables offers additional analytical insight. Comparison with the means reported for manipulators and non-manipulators in Beneish's original model (Table 4) may help identify which financial statement items may generate warning signals. However, higher variable values do not necessarily imply manipulation, as they may reflect normal business developments, industry-specific factors, or the stage of the business cycle. Additionally, since means may be substantially influenced by extreme observations, medians were examined, while sample standard deviations were used to describe dispersion within each group.

A two-sided Fisher's exact test was used to examine the association between the Altman and Beneish classifications, given its suitability for 2×2 contingency tables with low expected cell frequencies. The test has also been used in related research involving Altman models (Matanga & Holman, 2024) and financial statement manipulation detection models (Marais, Vermaak & Shewell, 2023). For this analysis, Altman classifications were grouped into two categories: within and outside the distress zone. The Beneish classification for year t was compared separately with the Altman classifications for years t and $t-1$. Statistical significance was assessed at the 5% level. The exploratory analysis assessed associations between model-generated classifications within the sample of bankrupt companies, rather than between financial distress and independently verified fraudulent financial reporting.

3. Data

The companies included in the sample were identified based on the list of active bankruptcy proceedings maintained by the Bankruptcy Supervision Agency in the period 2017–2024. Of the total of 3,320 companies in bankruptcy, 65 large and medium-sized companies were selected. The focus on these companies is justified by the economic and social consequences of their business failure, given the amount of capital employed, number of employees, turnover generated and the provision of activities of public importance (Miletić, 2019). Furthermore, these companies are expected to maintain a higher quality of financial reporting and more developed internal control systems, which should reduce the risk of accounting errors and fraud. Finally, large companies in Serbia apply International Financial Reporting Standards (IFRS), while medium-sized companies apply IFRS for SMEs, unless they are listed companies or parent companies with consolidated reporting obligations, or they voluntarily opt to apply the full set of IFRS, thereby contributing to the international comparability of financial results. Nevertheless, financial scandals and bankruptcies have often been associated with fraudulent financial reporting by large corporations worldwide (Beasley, Hermanson, Carcello & Neal, 2010), while in Serbia, according to the latest research by Malinić and Ribić (2025), more than 11% of large companies are classified as manipulators.

Graph 1 presents the distribution of the 65 selected companies by the year in which bankruptcy proceedings were initiated.

Graph 1: Initially selected bankrupt companies by year

Source: <https://alsu.gov.rs/>

The fewest bankruptcy proceedings were initiated in 2018, while the highest number was recorded in 2022. The 2022-2024 period accounts for more than half of the 65 initially selected companies undergoing bankruptcy proceedings, which may be associated with post-pandemic disruptions and geopolitical tensions. Their consequences were reflected in rising commodity and raw material prices, disruptions in supply chains, shortages and inventory accumulation, increased production costs, declining imports and exports, rising inflation, as well as disruptions in banking systems and capital markets (Pantelić & Živanović, 2024). The slight decline in the number of bankruptcy proceedings after 2022 may be associated with fiscal support measures, which may have temporarily delayed the formalization of some cases of insolvency. This dynamic may also have been the result of compensatory effects at the economy-wide level, as some sectors were more severely affected by crisis-related disruptions, while others operated under more favorable conditions (Malinić, 2023). Nevertheless, the consequences of bankruptcy remain serious, as the losses incurred by these companies are measured in millions.

In accordance with the Beneish model methodology, the analysis relies on data from balance sheets and income statements obtained from the Serbian Business Registers Agency website, for the year preceding the initiation of bankruptcy proceedings. Given that models based on financial variables are most commonly applied to predict bankruptcy one to two years before its occurrence (Altman, Iwanicz-Drozowska, Laitinen & Suvas, 2015), this criterion was adopted for the application of the Altman

model. Since financial statements for the observed period were unavailable for 17 companies, the sample was reduced to 48 companies. In addition, 8 companies were excluded from further analysis because the occurrence of zero values in the denominators of certain ratios made it impossible to calculate the Beneish indicators. The final sample comprises 40 companies, of which 80% are limited liability companies and 20% are joint-stock companies. In terms of sectoral distribution, manufacturing has the largest share (50%), followed by trade (18%), the primary sector (15%), construction (10%) and services (7%). The entire sample selection process is presented in Table 2.

Table 2: Sample formation process

Criterion	Number of companies
Total number of companies for which bankruptcy proceedings were initiated during the period 2017-2024	3,320
Medium-sized and large companies for which bankruptcy proceedings were initiated during the period 2017-2024	65
Companies for which basic financial statements were not available	(17)
Companies for which it was not possible to calculate the Beneish M-5 model variables	(8)
Final sample size	40

Source: Authors' calculations

The results of applying the Altman and Beneish models are presented, first through separate analyses and then through a comparative assessment of the findings.

4. Results and discussion

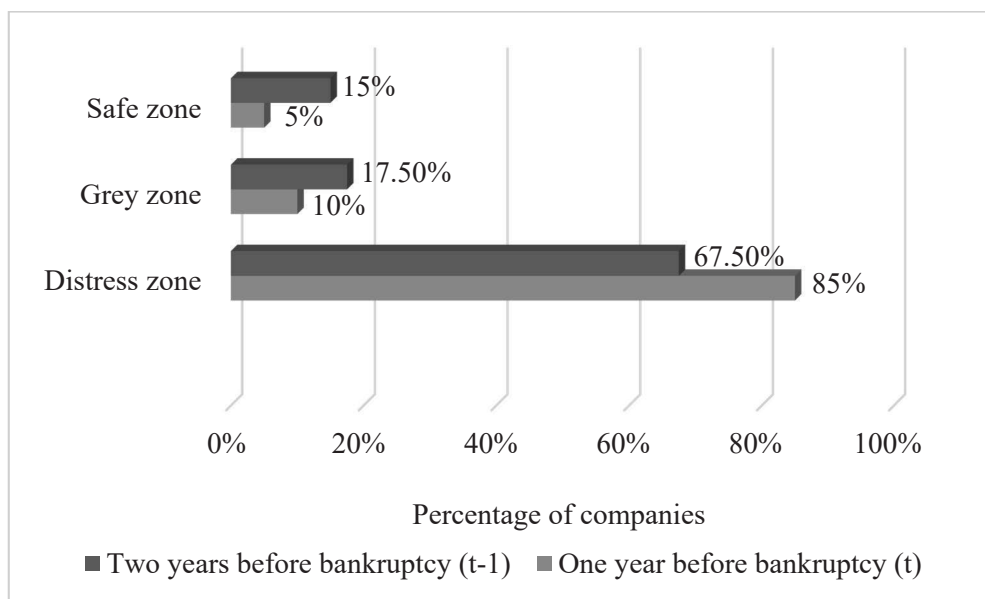
4.1. Application of the Altman model

Table 3 presents the descriptive statistics of the Altman Z"-score and the detection rates among the sampled bankrupt companies, while Graph 2 shows their distribution across the Altman risk zones. The results are reported for one year (t) and two years (t-1) prior to the initiation of bankruptcy proceedings. The detection rate, interpreted as sensitivity, represents the proportion of subsequently bankrupt companies classified in the distress zone ($Z'' < 1.10$). Since the sample contains only bankrupt companies, the detection rates do not represent overall classification accuracy.

Table 3: Summary statistics and bankruptcy detection rates of the Altman model

	One year before bankruptcy	Two years before bankruptcy
Mean	-11.13	-0.19
Standard deviation	36.24	3.19
Minimum	-183.78	-13.81
Maximum	12.13	4.50
Median	-2.87	0.21
Cut-off < 1.10		
Detection rate (sensitivity)	85.00%	67.50%

Source: Authors' illustration

Graph 2: Distribution of companies by Altman risk zones

Source: Authors' illustration

Both the mean and median Z"-score declined as bankruptcy approached, indicating a deterioration in the sampled companies' financial condition. The detection rate increased from 67.50% in year t-1 to 85.00% in year t, while the shares of companies in the *grey* and *safe* zones decreased. The low score values reflect the presence of several dimensions of financial risk: profitability risk (X_2 and X_3), solvency risk (X_4) and liquidity risk (X_1). This sequence follows the logic of the development of financial distress, which, when reaching a severe level, leads to bankruptcy.

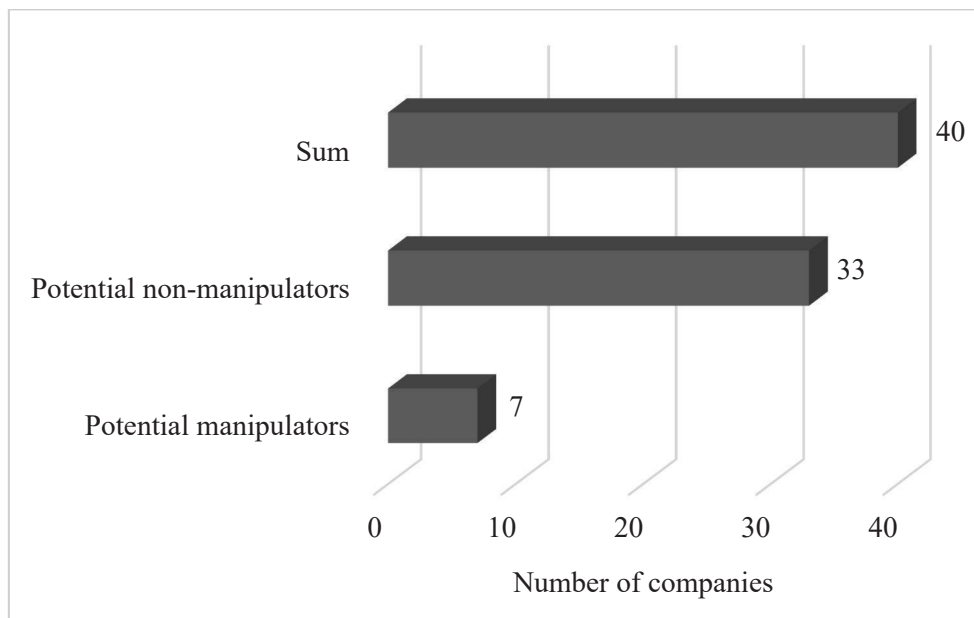
If the initial signs of crisis, which typically manifest themselves through a weakening of a company's strategic position, are not addressed in a timely manner, the consequences first become evident in profitability. This is reflected in the current

profitability indicators, whose values were close to zero or negative in almost all sampled companies. Moreover, the findings show that profitability was impaired not only in the current accounting period, but that the companies' operating problems also had a cumulative character. The accumulation of losses gradually eroded equity, thereby increasing companies' dependence on external financing. Solvency indicator values below one were recorded in 98% of the companies in both observed years. This means that, in almost the entire sample, equity was lower than total liabilities, i.e. that the companies relied to a substantially greater extent on external than on internal sources of financing. Among companies that still had positive equity, total liabilities were approximately 3.7 times higher than equity in year t-1 and approximately 3.9 times higher in year t, which, combined with the continuous accumulation of losses, further impaired their ability to service interest obligations. An even more unfavorable financial position was observed among a substantial proportion of the sampled companies, whose accumulated losses exceeded their total equity, resulting in equity being reported as zero in the balance sheet. This situation indicated that the companies were over-indebted and faced serious financial and structural imbalances, a high degree of long-term financial insecurity and threats to their continued operations in the foreseeable future. The decline in profitability and the increase in indebtedness over an extended period inevitably resulted in deteriorating liquidity, reflected in a decline in net liquid assets relative to total assets. This indicator was unfavorable in the majority of companies, as 60% had negative working capital in year t-1 and 80% of the sample in year t. There being no mechanisms to stabilize their ability to meet financial obligations, the companies ultimately became completely illiquid, which, together with over-indebtedness, constituted the legal basis for initiating bankruptcy proceedings.

The application of the Altman model should preferably be accompanied by an assessment of financial reporting quality, as *creative* accounting practices may undermine the reliability of the underlying financial data and, consequently, the model's results. Such an approach could enhance financial reporting transparency and strengthen the credibility of the accounting profession (Malinić, 2013).

4.2. Application of the Beneish model

The Beneish M-score model is used to identify indicators of financial statement manipulation. Given the limitations discussed below, its results should be interpreted as signals of possible fraudulent financial reporting rather than evidence of actual manipulation. As shown in Graph 3, 7 of the 40 companies (17.50%) were classified as potential manipulators and 33 (82.50%) as potential non-manipulators.

Graph 3: Distribution of companies by Beneish classification

Source: Authors' illustration

Table 4 presents the mean values of the individual variables and the overall M-score for both groups, alongside those reported in Beneish's study. The mean M-score is 4.40 for potential manipulators, well above the critical value, compared with -5.54 for potential non-manipulators. The individual variables exhibit a mixed pattern across the two groups, with the observed differences interpreted descriptively.

Table 4: Average values of variables and M-score for potential manipulators and potential non-manipulators

Indicator	Potential non-manipulators		Potential manipulators	
	Values for sampled companies	Beneish benchmark	Values for sampled companies	Beneish benchmark
DSRI	1.27	1.031	10.39	1.465
GMI	-2.05	1.014	0.21	1.193
AQI	1.12	1.039	2.24	1.254
SGI	0.69	1.134	0.42	1.607
DEPI	1.65	1.001	0.87	1.077
M-score	-5.54	< -2.22	4.40	> -2.22

Source: Authors' calculation

Table 5 presents the medians and sample standard deviations of the five Beneish variables, providing additional insight into their central tendency and dispersion

within each group. Group membership is based on the M-score classification, not on independently verified manipulation status.

Table 5: Medians and standard deviations of the Beneish variables by group

Indicator	Potential non-manipulators		Potential manipulators	
	Median	Standard deviation	Median	Standard deviation
DSRI	1.324	0.631	8.494	9.513
GMI	0.052	8.751	0.077	0.777
AQI	0.977	0.803	0.919	3.703
SGI	0.693	0.319	0.269	0.437
DEPI	0.943	3.075	0.927	0.261

Source: Authors' calculation

DSRI exhibits the most pronounced difference between potential manipulators and non-manipulators. Among potential non-manipulators, its mean slightly exceeds Beneish's benchmark, indicating an increase in receivables relative to revenues that may reflect collection difficulties amid financial distress. Among potential manipulators, both the mean and median are markedly higher, suggesting a broader divergence between receivables and revenues, similar to MacCarthy's (2017) findings. This may reflect not only the consequences of the crisis and collection difficulties but also the extension of credit periods granted to customers. Insufficient disclosure of changes in credit policies may reduce financial statement transparency and obscure the accumulation of overdue receivables. Related-party receivables warrant particular attention because the underlying transactions may deviate from market terms. Failure to recognize necessary write off uncollectible receivables may create an artificially favorable picture of a company's financial position and profitability. DSRI therefore serves as a warning indicator warranting closer examination of revenues, the structure of receivables, and customer credit policies.

High GMI values indicate deteriorating gross margins, which may increase incentives to manipulate reported results (Beneish, 1999). However, in our study, both groups had low, similar medians, while the negative mean among potential non-manipulators was influenced by extreme negative observations. The expected pattern was therefore not observed, consistent with Muria, Alim and Prasetyono (2024).

The mean AQI was higher among potential manipulators than among potential non-manipulators, although the medians were similar. Since the higher mean among potential manipulators was influenced by a single company, it should not be interpreted as a consistent pattern within this group. Nevertheless, elevated AQI values may warrant closer examination of items subject to managerial judgment,

such as intangible assets, including the accounting policies applied to their recognition and measurement, as well as related impairment testing.

Both mean and median SGI values were below 1 in both groups, indicating a general pattern of declining sales consistent with the companies' financial difficulties. This contrasts with the growth-related pattern discussed by Beneish (1999), Mahama (2015), and MacCarthy (2017), whereby sales growth may increase stakeholder expectations and pressure on management to sustain such growth.

High DEPI values may signal possible depreciation manipulation, as lower depreciation rates reduce expenses and increase reported profit (Beneish, 1999). In this sample, the higher mean among potential non-manipulators reflected extreme observations, while the medians were similar across the two groups. Elevated values may also reflect changes in asset composition or useful life estimates and do not, by themselves, establish intentional manipulation.

4.3. Comparative analysis of Altman and Beneish results

Since the Altman and Beneish models assess different types of risk, their comparison focuses on the coexistence and association between financial distress and potential manipulation signals rather than classification accuracy (Table 6). The Beneish classification for year t was first matched with the Altman classification for the same year and then with the classification for year $t-1$ to determine whether the companies classified as potential manipulators had also experienced financial distress one year earlier. The percentages were calculated as proportions of the total sample. Table 6 also reports the two-sided Fisher's exact test p -values for both comparisons.

Table 6: Integrated risk profile based on Altman and Beneish classifications

Combination of classifications	Number (t)	% (t)	Number ($t-1$)	% ($t-1$)
Distressed + potential manipulators	6	15.00%	6	15.00%
Distressed + potential non-manipulators	28	70.00%	21	52.50%
Non-distressed + potential manipulators	1	2.50%	1	2.50%
Non-distressed + potential non-manipulators	5	12.50%	12	30.00%
Total	40	100.00%	40	100.00%
Fisher's exact test (two-sided), p -value:	1.000		0.393	

Source: Authors' calculation

In year t , 28 companies (70%) were classified in the Altman distress zone without a Beneish manipulation signal, while 6 companies (15%) exhibited both signals. The

proportion of potential manipulators was similar within and outside the distress zone: 6/34 (17.65%) and 1/6 (16.67%), respectively. The two-sided Fisher's exact test indicated no statistically significant association between the classifications in year t ($p = 1.000$), providing insufficient evidence that companies in the distress zone were more likely to exhibit a Beneish manipulation signal. The absence of a Beneish signal does not, however, rule out financial reporting irregularities.

Retrospectively, 6 of the 7 companies classified as potential manipulators in year t were already in the Altman distress zone in year $t-1$. The proportion of potential manipulators was 6/27 (22.22%) among companies in the distress zone in year $t-1$ and 1/13 (7.69%) among those outside it, with no statistically significant association ($p = 0.393$). Among potential non-manipulators, the number of companies classified in the Altman distress zone increased from 21 in year $t-1$ to 28 in year t as bankruptcy approached.

5. Limitations of the research

The assessment of the Altman model in this study is limited by the inclusion of bankrupt companies only, without a control group of non-bankrupt companies. The reported detection rates indicate the proportion of subsequently bankrupt companies classified in the distress zone, but do not provide a measure of overall classification accuracy. Second, the choice of explanatory variables has been criticized because it is based on their prevalence in the literature rather than on robust theoretical foundations (Grice & Ingram, 2001). Third, Altman et al. (2015) claim that models based on financial variables can be used to predict bankruptcy one to two years before its occurrence, an approach also adopted in this study. However, analysis over a five-year period would make it possible to determine whether the risk had emerged earlier or increased substantially shortly before the initiation of bankruptcy proceedings. Fourth, the Altman model was developed several decades ago and is not fully adapted to contemporary business conditions (Boritz, Kennedy & Sun, 2007). Changes in attitudes toward risk, financing practices, prevailing levels of indebtedness and bankruptcy legislation call for periodic reassessment of the selected variables, their weights and cut-off values. Fifth, the observation period partly coincided with the pandemic and geopolitical tensions, which may have affected companies' financial indicators and their model-based classifications. The effects of these external circumstances were not separately examined in this study.

A limitation of the Beneish model is that its result does not constitute conclusive evidence of fraudulent financial reporting but merely indicates an increased probability of its occurrence. A thorough qualitative analysis of the results is necessary before drawing conclusions. The application of the Beneish model in the domestic context presents additional challenges, as differences in accounting practices may complicate the calculation and interpretation of its variables. Finally, the model was developed using data from publicly listed companies in the US, whereas Serbia's capital market is small and insufficiently liquid.

The small sample of 40 bankrupt companies, including only 7 classified as potential manipulators, limits statistical power and calls for caution when interpreting the non-significant Fisher's exact test results. The findings cannot be generalized to all Serbian companies or establish causality between financial distress and fraudulent financial reporting. Differences in Beneish variables between groups are descriptive and do not independently validate the classifications, as these variables were used to calculate the M-score and define the groups.

Conclusion

The application of the Altman Z"-score and Beneish M-score (M-5) models identified warning signals of financial distress and possible fraudulent financial reporting in the sampled Serbian companies prior to the initiation of bankruptcy proceedings. The proportion of companies classified in the Altman distress zone increased from 67.50% two years before bankruptcy to 85.00% one year before bankruptcy. The Beneish model classified 17.50% of the companies as potential manipulators, with DSRI showing the most pronounced difference between the two groups. Although the exploratory analysis using Fisher's exact test did not establish a statistically significant association between the two models' classifications, the coexistence of both warning signals in several companies should not be overlooked. Such cases may warrant closer examination, particularly because Beneish signals may help identify areas of financial reporting requiring further investigation.

Early identification of such patterns, combined with a broader assessment of financial difficulties, could support corrective action and consideration of out-of-court financial restructuring where feasible. Successful restructuring may enable the company to continue operating, preserve jobs and ownership interests, and improve creditors' prospects of recovering their claims, thereby mitigating some of the economic and social consequences of bankruptcy. Careful examination of suspected manipulative practices may also contribute to the credibility of the accounting profession and reinforce its role in promoting transparent and ethical reporting. If bankruptcy proceedings are unavoidable, the detection of fraudulent financial reporting may assist investigative authorities in establishing responsibility and pursuing appropriate legal action against individuals involved in unlawful activities.

Differences among national economies in terms of economic structure, access to capital, risk attitudes among companies, reporting culture, the level of development of accounting practices and the quality of institutional oversight highlight the importance of assessing model suitability across national contexts. Future research should therefore evaluate and adapt existing models or develop new ones for the Serbian context, using larger samples that include both bankrupt and non-bankrupt companies. Such models could enhance the ability of accountants, auditors and investigators to identify high-risk patterns in a timely manner, thereby helping to mitigate the adverse consequences associated with financial crises and fraudulent financial reporting.

Acknowledgement:

For the first author (Jovana Stojanović): This research was financially supported by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia (Contract No. 451-03-33/2026-03/200371).

For the second author (Dejan Spasić): The paper is the result of research based on the commitments made under the Funds Transfer Agreement for the financing of NIR in 2026 (registration number 451-03-34/2026-03), concluded between the Ministry of Science, Technological Development and Innovation of the Republic of Serbia and the Faculty of Economics of the University of Niš.

References

- Altman, E. I., Hartzell, J., & Peck, M. (1995). *Emerging markets corporate bonds: A scoring system*. New York: Salomon Brothers Inc.
- Altman, E. I., Iwanicz-Drozowska, M., Laitinen, E. K., & Suvas, A. (2015). Financial and non-financial variables as long-horizon predictors of bankruptcy. *SSRN 2669668*. <https://doi.org/10.2139/ssrn.2669668>
- Association of Certified Fraud Examiners. (2026). *Occupational Fraud 2026: A Report to the Nations*, retrieved on September 1, 2026, from: <https://www.acfe.com/acfe-insights-blog/blog-detail?s=key-findings-report-to-the-nations-2026>
- Beasley, M. S., Hermanson, D. R., Carcello, J. V., & Neal, T. L. (2010). Fraudulent financial reporting: 1998-2007: An analysis of US public companies. *Association Sections, Divisions, Boards, Teams*. 453.
- Beneish, M. D. (1999). The detection of earnings manipulation. *Financial Analysts Journal*, 55(5), 24-36. <https://doi.org/10.2469/faj.v55.n5.2296>
- Beneish, M. D., Lee, C. M., & Nichols, D. C. (2013). Earnings manipulation and expected returns. *Financial Analysts Journal*, 69(2), 57-82. <https://doi.org/10.2469/faj.v69.n2.1>
- Boritz, J., Kennedy, D., & Sun, J. (2007). Predicting business failures in Canada. *Accounting Perspectives*, 6(2), 141-165. <https://doi.org/10.1506/G8T2-K05V-1850-52U4>
- Carton, R. B., & Hofer, C. W. (2006). *Measuring organizational performance: Metrics for entrepreneurship and strategic management research*. Edward Elgar Publishing. <https://doi.org/10.4337/9781847202840>
- Chan, C., & Landry, S. P. (2019). Financial statements too good to be true? An instructional case assessing that question using analytical procedures and Beneish's M-score. *Journal of Forensic and Investigative Accounting*, 11(2), 380-394.

Compagnino, A. A., Maruccia, Y., Cavuoti, S., Riccio, G., Tutone, A., Crupi, R., & Pagliaro, A. (2025). An introduction to machine learning methods for fraud detection. *Applied Sciences*, 15(21), 11787. <https://doi.org/10.3390/app152111787>

Deloitte Forensic Center. (2008). *Ten things about bankruptcy and fraud: A review of bankruptcy filings*. Deloitte Financial Advisory Services LLP.

Dimitrijević, D. S., Obradović, V. M., & Milutinović, S. (2018). Indicators of fraud in financial reporting in the Republic of Serbia. *TEME*, 42(4), 1319-1338. <https://doi.org/10.22190/TEME1804319D>

Dimitrijević, D., Milutinović, S., & Stanković, P. (2024). Determinants of the occurrence of financial distress in medium-sized and big public joint-stock companies. *Economic Horizons/Ekonoski Horizonti*, 26(3), 301-319. <https://doi.org/10.5937/ekonhor2403301D>

Dyck, A., Morse, A., & Zingales, L. (2024). How pervasive is corporate fraud?. *Review of Accounting Studies*, 29(1), 736-769. <https://doi.org/10.1007/s11142-022-09738-5>

Grice, J. S., & Ingram, R. W. (2001). Tests of the generalizability of Altman's bankruptcy prediction model. *Journal of Business Research*, 54(1), 53-61. [https://doi.org/10.1016/S0148-2963\(00\)00126-0](https://doi.org/10.1016/S0148-2963(00)00126-0)

MacCarthy, J. (2017). Using Altman Z-score and Beneish M-score models to detect financial fraud and corporate failure: A case study of Enron Corporation. *International Journal of Finance and Accounting*, 6(6), 159-166. <https://doi.org/10.5923/j.ijfa.20170606.01>

Mahama, M. (2015). Detecting corporate fraud and financial distress using the Altman and Beneish models. *International Journal of Economics, Commerce and Management*, 3(1), 1-18.

Malinić, D. (2013). Merenje performansi primenom scoring modela u cilju otkrivanja ranih upozoravajućih signala [Performance measurement using scoring models to detect early warning signals]. U: Mikerović, D. (Ured.), *Računovodstvo, revizija i poslovne finansije u uslovima odgovornijeg ponašanja učesnika na tržištu*, (str. 95-120). Banja Luka: Savez računovođa i revizora Republike Srpske.

Malinić, D. (2016). Mogućnosti i rizici ocene performansi privrede Srbije na osnovu analize finansijskih izveštaja [Opportunities and risks in assessing the performance of the Serbian economy based on financial statement analysis]. *Računovodstvo*, 1, 7-37.

Malinić, D. (2023). Posledice pandemije COVID-19 na performanse privrede Republike Srbije [Consequences of COVID-19 pandemic on the performance of the economy of Serbia]. U: Dimitrijević, D., Lalević Filipović, A., & Spasić, D. (Ured.), *Računovodstvena znanja kao činilac ekonomskog i društvenog napretka* (str. 3-19). Kragujevac: Univerzitet u Kragujevcu, Ekonomski fakultet.

Malinić, D., & Ribić, M. (2025). Identifying indicators of fraudulent financial reporting under conditions of institutional opacity: The relevance of the Beneish model. In: Malinić, D., & Vučković Milutinović, S. (Eds.), *Contemporary Forensic Accounting Trends in Combating Financial Fraud: Global and National Perspectives* (pp. 1-22). Belgrade: University of Belgrade, Faculty of Economics and Business.

Marais, A., Vermaak, C., & Shewell, P. (2023). Predicting financial statement manipulation in South Africa: A comparison of the Beneish and Dechow models. *Cogent Economics & Finance*, 11(1), 2190215. <https://doi.org/10.1080/23322039.2023.2190215>

Matanga, N., & Holman, G. (2024). Adapting Altman Z-score models for early warning signals: Evidence from delisted mining stocks on the Johannesburg Stock Exchange. *Investment Analysts Journal*, 53(3), 249-261. <https://doi.org/10.1080/10293523.2024.2397892>

Miletić, S. (2019). Savremeni pristupi merenju kvaliteta finansijskih izveštaja [Contemporary approaches to measuring the quality of financial statements]. U: Stojanović, R., Škobić, Z., Nešić, V., Petrović, P., Kilibarda, Z. (Ured.), *Finansijsko poslovanje i izveštavanje u funkciji digitalne ekonomije i održivog razvoja* (str. 246-265). Beograd: Savez računovođa i revizora Srbije.

Muminović, S., Pavlović, V., & Cvijanović, J. M. (2012). The impact of investments and changes in the production regime on the results of creditworthiness assessment and bankruptcy prediction models: Case study: Company Bulgari Filati d.o.o. *Industrija*, 40(2), 3-18.

Muñoz-Izquierdo, N., Camacho-Miñano, M. D. M., & Pascual-Ezama, D. (2017). Contenido del informe de auditoría en el año previo a la declaración del concurso de acreedores. Contraste empírico para el caso español [The content of the audit report in the year prior to bankruptcy filing. Empirical evidence from Spain]. *Spanish Journal of Finance and Accounting/Revista Española de Financiación y Contabilidad*, 46(1), 92-126. <https://doi.org/10.1080/02102412.2016.1245886>

Muñoz-Izquierdo, N., Laitinen, E. K., Camacho-Miñano, M. D. M., & Pascual-Ezama, D. (2020). Does audit report information improve financial distress prediction over Altman's traditional Z-Score model?. *Journal of International Financial Management & Accounting*, 31(1), 65-97. <https://doi.org/10.1111/jifm.12110>

Muria, R. M., Alim, M. N., & Prasetyono. (2024). The ability of the Beneish M-score to detect the trends of fraud in the Indonesian Sharia Stock Index. *Journal of Economics, Finance and Management Studies*, 7(1), 756-770. <https://doi.org/10.47191/jefms/v7-i1-76>

Özari, Ç., Can, E. N., & Demirkale, Ö. (2025). Financial fraud detection with Altman Z-score and Beneish M-score via random forest: Verified by Borsa Istanbul fines (2018–2022). *SAGE Open*, 15(4). <https://doi.org/10.1177/21582440251386174>

Pantelić, M., & Živanović, M. (2024). *Finansijsko izveštavanje [Financial reporting]*. Beograd: Univerzitet u Beogradu, Ekonomski fakultet.

Schilit, H. M., Perler, J., & Engelhart, Y. (2018). *Financial shenanigans: How to detect accounting gimmicks and fraud in financial reports*. New York: McGraw-Hill.

Singh, B. P., & Mishra, A. K. (2016). Re-estimation and comparisons of alternative accounting based bankruptcy prediction models for Indian companies. *Financial Innovation*, 2(1), 6. <https://doi.org/10.1186/s40854-016-0026-9>

Škarić-Jovanović, K., & Spasić, D. (2022). *Specijalni bilansi [Special-purpose financial statements]*. Beograd: Univerzitet u Beogradu, Ekonomski fakultet.

Soukal, I., Mačí, J., Trnková, G., Svobodova, L., Hedvičáková, M., Hamplova, E., Maresova, P., & Lefley, F. (2024). A state-of-the-art appraisal of bankruptcy prediction models focussing on the field's core authors: 2010–2022. *Central European Management Journal*, 32(1), 3–30. <https://doi.org/10.1108/CEMJ-08-2022-0095>

Spasić, D., & Stojanović, J. (2026). The nexus between fraudulent financial reporting and company failure: The case of bankrupt companies in Serbia. *TEME*, L(2), 437–460. <https://doi.org/10.22190/TEME260131028S>

Svoboda, J., Berková, I., Pražáková, J., & Vejsadová Dryjová, M. (2025). Relationships between the Beneish M-score and bankruptcy models: Insights from financial models in Czech manufacturing firms. *Proceedings of Rijeka Faculty of Economics: Journal of Economics and Business*, 43(2), 465–484. <https://doi.org/10.18045/zbefri.2025.2.7>

Velte, P. (2023). The link between corporate governance and corporate financial misconduct. A review of archival studies and implications for future research. *Management Review Quarterly*, 73(1), 353–411. <https://doi.org/10.1007/s11301-021-00244-7>

Vučković Milutinović, S. (2019). Analysis of modifications to auditor's opinion on financial statements of listed companies in Serbia. *Ekonomika preduzeća*, 67(3–4), 212–223. <https://doi.org/10.5937/EKOPRE1904212V>

Wang, M., Zhang, X., & Han, X. (2025). AI driven systems for improving accounting accuracy fraud detection and financial transparency. *Frontiers in Artificial Intelligence Research*, 2(3), 403–421. <https://doi.org/10.71465/fair398>

<https://alsu.gov.rs/>

<https://apr.gov.rs>

Mario Abela
John Jay College of Criminal Justice,
City University of New York
ORCID: 0000-0003-1427-7740

ANATOMY OF A SCANDAL: DWS FROM AN INSTITUTIONAL LOGICS PERSPECTIVE

UDK 343.53: 502.131.1
DOI <https://doi.org/10.54318/czf.2026.ma.014>

Abstract

Greenwashing has been rife in investment products claiming to meet sustainable investment criteria but later exposed as being misleading and fraudulent. More recently, it has attracted the attention of regulators both in terms of legislation and enforcement action. The literature often employs the fraud triangle to examine the circumstances under which individuals perpetrate fraud. To ameliorate this risk, governance requirements place emphasis on the importance of internal controls and the control environment without much regard to the incentives for overriding them when the norms and beliefs of the leadership of the organization sanction that behavior.

A wider view is taken, employing institutional theory to examine the decoupling between the ‘talk’ and the ‘walk’, breaking that down into competing institutional logics that justify breaches of internal controls and governance practices. It reveals the existence of a dominant market logic endorsing bad behavior in the interests of pursuing short-term financial gains. The DWS scandal is analyzed to demonstrate the dynamics of institutional-wide greenwashing. The analysis signals that what constitutes ‘fraud’ is context and time specific and ambiguities in the how market participants construct and assess the validity of green claims is exploited until regulators set appropriate guardrails and enforcement actions. Relying on the agency of whistleblowers within an organization puts those individuals in a precarious situation with often dire consequences.

This paper contributes to the literature by applying an institutional logics perspective to analyze fraud in a major financial institution, highlighting that laws and regulation that rely heavily on individual judgement are ineffective policy instruments if they fail to adequately address the incentives system within the financial institution that they attempt to regulate.

Key words: *sustainable finance, greenwashing, fraud, regulation, institutional theory*

JEL Classification: *M41, Q56, K13*

Introduction

The prevalence of greenwashing, notably making false claims about the sustainability credentials of investment products, has become a prominent issue for the capital markets (ESMA, 2024). In most cases ‘greenwashing’ is fraud as it represents a misrepresentation of the facts (Abela, 2025). Legislation, investment taxonomies and enforcement actions have been some of the tools that regulators have used to combat misleading practices. For investors, a recent European survey found that misleading, hard-to-verify and the unreliable nature of green claims, was cited as the main reason that they would not invest in sustainable products (BEUC, 2025, p. 7). Despite the urgent and well-documented need to respond to the environmental crisis of our times, these practices have served to further stall the transition of capital toward more sustainable business models. As Perez & Verbruggen (2026a) observe, leaving it to the market and private contracts is resulting in ‘green defaults’, where so-called ‘sustainable’ financial instruments fail to meet their sustainability targets, creating ‘green junk’ investment products. It has consequently led to a degree of cynicism and a lack of trust in what is labelled as a ‘green’ investment product (BEUC, 2025).

The literature has focused on the application of the fraud triangle as a basis for analyzing cases of greenwashing by companies and in sustainable finance products (Kurpierz & Smith, 2020; Marquis et al., 2016; Yu & Hwang, 2026). The application of the fraud triangle to greenwashing has coined the term the “greenwashing triangle” to denote a variant to the traditional notion of fraud in the literature (e.g. Kurpierz & Smith, 2020). This paper takes a broader view, to consider greenwashing in investment products at the institutional level, applying institutional theory to analyze the dynamics of organizational governance and change. In that respect it follows a similar analytical approach to Hartmann et al. (2018) who recognizes the need for more case studies to investigate fraud. Furthermore, they note that research on fraud “needs contextualization and call for further research on its social construction, temporality and cultural embeddedness” (Hartmann et al., 2018, p. 847). It has long been recognized in the organizational literature that institutions often separate their formal structures or what is documented in policies and day-to-day practices (Meyer & Rowan, 1977) sometimes referred to as the difference between the ‘talk’ and the ‘walk’ to signify the decoupling that occurs in institutions. This case study of the scandal at Deutsche Gesellschaft für Wertpapiersparen (DWS) establishes *par excellence* the dynamics of decoupling that led to widespread fraud in a major German financial institution. The DWS Scandal costing the company \$50m in fines, the largest ever imposed for fraud relating to greenwashing, from the German and US regulators (Palma & Storbeck, 2023) illustrates how a dominant institutional market logic (Thornton & Ocasio, 2008) legitimized behaviors that ignored compliance and regulatory logic to sustain an institution-wide fraud that if not for a whistleblower, Desiree Fixler, may have remained unchecked for many years (Fixler & Eccles, 2025). Based on publicly available interviews and regulatory documents, this paper presents a case study of the DWS scandal, analyzed through

an institutional logics perspective, to understand how the governance system and its internal controls were bypassed for a major financial institution in one of the world's most regulated markets.

The paper contributes to the accounting literature by demonstrating how well-accepted policies and practices that make an effective control environment can effectively be 'switch off' under the watchful eyes of management and the board. To the author's knowledge this is the first case study to employ an institutional logics perspective to the DWS case as a basis for examining the dynamics at play in such a fraud. Like many scandals, it should not have happened as the checks and balances within the corporation should have identified and prevented it for happening. The paper proceeds as follows: Section 1 provides the methodology used, Section 2 follows with a detailed examination of the available evidence and relevant theoretical framing, Section 3 draws policy implications and finally the conclusion signals the challenges of undertaking such a study and what lessons can be learned from it.

1. A brief note on methodology

A common challenge of analyzing fraud scandals is access to evidence, especially where the corporation continues to operate and dominate the narrative. The use of case studies as means for developing a detailed understanding of the dynamics within an organization is commonly used in the social sciences (Yin, 1994). As with any qualitative method there are attendant limitations or caveats that need to accompany the use of a case as a basis for understanding phenomenon. As Flyvbjerg (2006) puts it, the narrative of the event is the compelling element and the case study is a technique to elucidate that story: "Good social science is problem driven and not methodology driven in the sense that it employs those methods that for a given problematic, best help answer the research questions at hand" (p.242). Generalizability comes from understanding the distinct features and circumstances of multiple cases (Kuhn, 1987).

The legitimacy of a case study, and its suitability for addressing the research question, derives from the triangulation of multiple sources of evidence to cross-check details and accounts of events (Tight, 2010). In the present case, little direct evidence exists regarding internal events at DWS. This study therefore relied on publicly available documents and records, identified through extensive online searches conducted in July and August 2026 and covering the period from January 2021 to August 2026. The sources consulted included academic databases (EBSCO), newspaper coverage (e.g., the *Financial Times*), multimedia content (e.g., interviews on YouTube), and official public records, such as official documents on the website of the U.S. Securities and Exchange Commission (SEC). Search terms such as "DWS," "Deutsche Bank," "greenwashing," and "ESG" were combined to locate relevant material. Generative artificial intelligence (Claude Sonnet 5; Anthropic, 2026) was also used to identify additional sources and to check the completeness of the search; all sources suggested by the tool were independently verified against the

original documents before inclusion. All retrieved material was reviewed, and all items relating to DWS and ESG were analyzed and informed the findings of the research. (Tight, 2010). Accordingly, this case study draws on the evidence including articles and recorded interviews with Desiree Fixler who was the head of sustainability at DWS and the whistleblower that uncovered the fraudulent practices about the Environmental, Social and Governance (ESG) credentials of their investment products. Taking a constructivist perspective (Berger & Luckmann, 1991), understanding of events and practices is a product of the individuals own sensemaking (Brown et al., 2015). As Weick (1995) notes sensemaking is less about discovery and more about what people generate when interpret events. That is:

People engage in partially overlapping processes in which they construct ‘realities’ and then retrospectively make sense of them in a continuing dialogue of discovery and invention in which identities and social worlds are concomitantly referenced and fabricated. (Brown et al., 2015, p. 267)

It is against this context that the testimony of Desiree Fixler is examined. The use of evidence of whistleblowers is not unique in the examination of financial fraud (Shy, 2026) though much of the literature is focused on the individual rather than the institutional environment and prevailing institutional logics. Content analysis of DWS annual reports, included word searches to identify claims about sustainability and ESG. Different sections within the annual report were read and reviewed to better understand the narrative DWS constructed to explain their performance particularly around ESG (Krippendorff, 2013). As both proceedings initiated by the German Prosecutor and the SEC were settled prior to trial, there are no trial documents to examine, and details of the fraud and investigations carried out have not been made public.

2. Unpacking the DWS Scandal

2.1. Providing an Account

As noted by Hartmann et al. (2018), accounting, law and economics all have different concepts of “fraud”. The concept is challenging to anchor as the distinction between what is a matter of professional judgement and strict legal compliance remains problematic. This is especially true when it comes to sustainability, where the meaning of the term itself and related practices are so diverse that misrepresentations are constructed as marketing discretion rather than fraud. Can and Turker (2025) refer to this as the gray zone driven by external pressures to be ‘green’. However, the lack of clarity remains across different fields:

The voluntary nature of prevailing governance frameworks, combined with persistent doctrinal barriers and renewed political resistance to ESG initiatives, casts doubts on whether the interaction between regulatory and private enforcement mechanisms will be sufficient to shift market practices

towards more credible, transparent, and enforceable sustainability commitments. (Perez & Verbruggen, 2026b, p. 194)

What has become evident from the DWS case is that ambiguity was exploited for commercial gain. As the investment arm of Deutsche Bank, it was found to have misrepresented its investment products and the degree to which ESG was integrated into them. It was through the actions of a whistleblower, Desiree Fixler, an executive and head of sustainability at DWS, that it was revealed that the claims in their annual report were false and misrepresented the ‘greenness’ of investment products being sold. Fixler’s employment was terminated on 1 March 2021 after making accusations that the 2020 DWS Annual Report misrepresented its investments by stating that over 50% (\$540 billion) had been through a process it called “ESG Integration” (Kowsmann & Brown, 2021). As a significant example of how the policy was not applied, Fixler cites the major investment in Wirecard DWS maintained even after the company collapsed and the CEO was arrested (2023). There are 275 references to sustainability in the Annual Report and the topic features prominently in the letter to shareholders:

As a firm, we have placed ESG at the heart of everything that we do...our collective desire to help the communities in which we operate is embedded in our DNA, both as a fiduciary and as a corporate. And having previously identified ESG as one of the key mega-trends shaping the next decade – an assessment further underscored by the pandemic – we worked intensively to implement a comprehensive ESG strategy as we see it becoming a dominant theme for our investors, clients and regulators alike. We established a Group Sustainability Office to oversee a coherent and holistic ESG strategy firm-wide. We secured the support of high-calibre external experts for our new ESG Advisory Board. And we enhanced both our ESG integration as well as our stewardship efforts with the introduction of Smart Integration into the investment process. (DWS Group, 2021, p. 7)

This letter is signed by the Executive Board and proceeded by similar sentiments expressed by the Supervisory Board. What is remarkable, is that these statements and many like them, are contained in the legal financial filing. This is not a marketing document but one of record, that comes with strict legal obligations to present fairly the performance and position of the entity, especially given the changes to the German Commercial Code in the last decade (Hellmann et al., 2013).

The following statement from the Management Report in the Annual Report claims that the entity has a robust and comprehensive system of internal controls applicable to both financial and non-financial data:

In 2020, to support the accurate reporting of material non-financial information the Group has developed a non-financial internal control system, distinct from the internal control system for financial reporting. The objective of the internal control system for the integrated non-financial group statement is to provide accurate reporting of internally compiled

sustainability KPIs, as disclosed in the ‘Our Strategy - Sustainability Strategy’ section of this report. (DWS Group, 2021, p. 72)

There is no escaping that these statements were knowingly made based on having governance processes in place to test and evaluate the efficacy of these statements and claims.

In an interview with the *Financial Times*, Fixler recounts telling the board that she had concerns about the seriousness of misstatements being made in the Annual Report about ESG. Fixler said the CEO Asoka Wöhrmann stopped the meeting and he sent around a slide to the board and to her – the slide showed a collection of quotes from reports by sell-side analysts at other banks and he turned to Fixler and allegedly said “look the market believes we are a leader. Why don’t you?” (Tindera & Temple-West, 2022). Fixler was dismissed and claimed compensation for unfair dismissal but lost her case (Matussek & Arons, 2022). The management of DWS launched a probe by PwC to investigate the allegations made by Fixler. In May 2021 PwC issued a 42-page report that found that it “could not identify substance within the allegations” made. It led to DWS making a statement that it “stands by its disclosures in its annual reports. DWS firmly rejects the unfounded allegation being made by a former employee”.

Fixler’s comments led to investigations by the Securities and Exchange Commission (in the United States of America) and Bafin (German Federal Financial Supervisory Authority), with both taking action and fining DWS.

The SEC stated in its Order against DWS Investment Management Americas (DIMA):

Despite the statements made in its Annual Report and website, DWS was found to have misled investors and the SEC levied a fine of \$19m stating that:

DIMA made materially misleading statements about its controls for incorporating ESG factors into research and investment recommendations for ESG integrated products, including certain actively managed mutual funds and separately managed accounts. The order finds that DIMA marketed itself as a leader in ESG that adhered to specific policies for integrating ESG considerations into its investments; however, from August 2018 until late 2021, DIMA failed to adequately implement certain provisions of its global ESG integration policy as it had led clients and investors to believe it would. The order finds that DIMA also failed to adopt and implement policies and procedures reasonably designed to ensure that its public statements about the ESG integrated products were accurate. (Securities Exchange Commission, 2025)

In response, DWS stated:

The SEC ESG Order, following an extensive two-year examination, finds no misstatements in relation to our financial disclosures or in the

prospectuses of our funds. We have consistently stated that we stand by our financial disclosures and disclosures in our fund prospectuses.

The Order also makes clear that there was no intent to defraud, and the weaknesses identified by the SEC are in relation to processes and procedures that the firm has already taken steps to address. The Order also does not find that DWS Investment Management Americas (DIMA) staff were not integrating ESG factors into the investment process. Instead, the Order focuses on the fact that DIMA lacked processes and procedures to evidence that such integration was documented in a consistent manner.

We are pleased that the SEC recognized our cooperation in the investigation and our remediation efforts.

We have learned many lessons from the ever-evolving regulatory environment and are fully committed to continuing to improve. We are grateful to our clients for their continued trust in our products. (DWS Group, 2023)

The enforcement action and resulting fines were the largest ever given for greenwashing (Müller & Storbeck, 2025; Palma & Storbeck, 2023). It is interesting that in the 2025 Annual Report there are 590 references to sustainability but no mention of it in the letters to shareholders or in claims about its products. Most of the references to sustainability relate to compliance with CSRD and ESRS requirements. It remains curious why KPMG, as statutory auditor of DWS, was not subject to any enforcement action by the SEC or German prosecutors, and it is not mentioned in any of the official publicly available documents.

2.2. Applying Institutional Theory

The accounting perspective on fraud continues to be heavily influenced by an individual perspective, emphasizing that individuals have agency and therefore are free to act. The notion that the action is the product of a ‘script’ is an anathema to traditional fraud analysis. Furthermore, that those ‘scripts’ are rooted in ideologies and perceptions of rights and human intervention do not ring true with the dominant agency theory that continues to underpin accounting practice and control systems. Institutions, including financial ones, serve to constrain but also empower actions by actors (Jepperson & Meyer, 2021). Meyer and Rowan (1977) noted that formal structures and rules are often decoupled from actual activities. Despite organizations having formal policies and procedures, how individuals act within that environment is often different, being observed as dissonance between the ‘talk’ (what we say we do) and the ‘walk’ (how we actually behave). This phenomenon of decoupling has been applied extensively to sustainability particularly in providing a theoretical anchor to explain the prevalence of greenwashing (e.g. Jastram & Foersterling, 2024; Kauppi & Hannibal, 2017). In the case of DWS, decoupling is evident in the extensive policies and claims about integrating sustainability in investment products and the reality which was far from that. In moving from the individual level to the

organizational one, emphasis shifts from individual actors and their actions to giving accounts of how changed distributions in social values and opportunities occur. In the case of DWS, the analysis pivots from the perspective of the whistleblower to wider constructs within the organization that facilitated a dominant logic to override others and become normalized.

Institutional orders give rise to specific institutional logics. There are four core principles to an institutional logics perspective: the duality of agency and structure; institutions as material and symbolic, institutions as historical contingent and institutions as multiple levels of analysis. The interests, identities values and assumptions of individuals and organizations coexist in the institutional logics (Thornton et al., 2012). The idea of an institutional logics perspective has taken hold of organizational studies and led to the proliferation of all sorts of patterns of behavior to be labelled a type of 'logic'.

'Institutional logics' is defined as a basis for:

Analyzing the interrelationships among institutions, individuals, and organizations in social systems. It aids researchers interested in questions of how individual and organizational actors are influenced by their situation in multiple social locations in an interinstitutional system, for example the institutional orders of the family, religion, state, market, professions and corporation. (Thornton et al., 2012, p. 2)

One strand of the institutional logics literature focuses on conflicting institutional logics within an organization. Various approaches to understanding conflicts involve hybridization and selective coupling although other mechanisms have been identified such as deferent behavior, geographical separation and symbolic interaction. As Durand and Thornton highlight:

What matters in these different views of institutional change is that actors at any level have some capability to adjust their reference to the dominant institutional logic, therefore affecting its valence in the market or field and opening an alternative for conformity to and deviance from it. (2018, p. 635)

In the case of whistleblowers, they are in the difficult and precarious position of destabilizing the dominant institutional logic. As a result the pervasive bounded rationality of other actors moves into play to swiftly stabilize the organization (Besharov & Smith, 2014).

3. Analysis

In the many interviews and presentations Fixler has made about her experience at DWS, the key accusation she points to is a case of decoupling (Meyer & Rowan, 1977): "a huge gap between the 'say' and the 'do'" highlighting that there was no tracking system at DWS to support the claim that "the majority of assets under management were ESG compliant" (Fixler, 2023). As Fixler noted in an interview

“Yes, it is true that the ESG market lacks clear definitions, reporting standards and impact measurement guidelines. That is why greenwashing is so pervasive” (Bartz, 2021). She also recognized the complex nature of enforcing policies and acknowledge that “at any large company someone’s doing something stupid” (Fixler, 2023).

The employment of institutional logics tends to arise where actors face uncertainty in the actions they should take, with divergent views on how behaviors and motives should be interpreted (McPherson & Sauder, 2013). It highlights the difficulty of arriving at a universally shared concept of fraud and is underscored by the challenges of locating a common approach to the interpretation and compliance with standards and regulations (Hartmann et al., 2018). There are often no ‘bright-lines’ when it comes to applying sustainability standards which has made it difficult for regulators to enforce them (European Parliament, 2020). As Friedland and Alford (2013) note:

Institutions constrain not only the ends to which their behavior should be directed, but means by which those ends are achieved. They provide individuals with vocabularies of motives and with a sense of self. They generate not only what is valued, but the rules by which it is calibrated and distributed. Institutions set the limits on the very nature of rationality and, by implication, of individuality. Nonetheless, individuals, groups, and organizations try to use institutional orders to their own advantage (p. 251).

As a counter narrative to Fixler’s accusation, DWS argues “as we are already late to the game, we need to start the transformation” (Kowsmann & Brown, 2021), thereby endorsing its bullish approach to ‘greenness’ of its products. Setting an ambition is often invoked as the basis for justifying over-claiming sustainability actions. Greenwashing, as illustrated in the DWS case, “with media presentations, press releases, prospectuses, annual reports...was telling the markets that it was quite advanced when it comes to ESG investing” (Fixler, 2023). According to the rules set by DWS, it considered Fixler’s position to be untenable and at odds with the market logic it employed. It was noted in the Wall Street Journal, “DWS’s experience underscores the difficulties and pressure money managers are facing to plant their flag in the hottest corner of the fund market” (Kowsmann & Brown, 2021). It is that market signal that DWS leadership was attempting to respond to in order to generate value for shareholders, which it placed, according to the regulators, above the need to have effective controls and systems in place (Securities Exchange Commission, 2025).

Fixler presented a professional and ethical logic driven by what she described as “standing up for what’s right” (Bartz, 2021). Fixler notes emphatically that “not only do we not break the law but we shouldn’t break the spirit of the law, don’t get cute, don’t get fancy, don’t spin, don’t mis-sell, don’t misrepresent and don’t lie to your shareholders” (Fixler, 2023). It could be contended that Fixler’s statement is equally aspirational and subjective to implement in practice. Accordingly, logics are employed in this case to reinforce the perspective and logic imperative of the actor.

They, in turn give rise to practices that normalize and legitimize certain ways of working (Mutch, 2019).

Logics are not universal truths but constructions that develop and take shape over time so in that sense they are historically contingent (Friedland & Alford, 2013). That is important not only in framing what constitutes ‘fraud’ but also in the invocation of values and beliefs to support institutional decision-making. In that regard, locating the DWS into its historical moment is critical: “we need to think not just about moments of institutional selection and moments of institutional change, but of processes of institutional development unfolding over significant periods of time” (Pierson, 2004, p. 15). The DWS scandal occurred at a moment when policymakers recognized that the regulatory framework for sustainable finance was lacking and needed significant guardrails to ensure that corporations change their practices towards a more sustainable future (European Commission, 2019). It followed the European Commission’s own assessment that the Non-financial Reporting Directive was far too loose and failed to bring about the meaningful sustainability reporting supported by evidence of change (European Parliament, 2020). The existence of multiple logics or competing and conflicting logics is a common dynamic in organizations (Thornton et al., 2012).

In the case of DWS, balancing market pressures with compliance requirements puts into stark relief the trade-offs and consequences of the sensemaking by its leadership. At play is a dominant logic of extraction and market advantage above all else, and that logic captures artefacts and rituals within the organization to establish a de facto set of rules that stated policies and procedures can be ignored, as it encodes a criterion of legitimacy. The arguments put forward by the DWS leadership rested heavily on a degree of isomorphic mimicry (Dimaggio & Powell, 1983) – ‘everyone else is doing it’. The SEC in its assessment of practices at DWS notes:

research analysts having inconsistent levels of documented compliance with the ESG Integration Policy’s requirements to consider material ESG risk factors in research and valuation models. The ESG Integration Policy nonetheless remained published on the website...[when it] knew or should have known that it lacked adequate procedures to ensure this was the case...(SEC Order, 2023, p. 2)

It is unsurprising that DWS settled this complaint without challenging the SEC’s findings because as Fixler had recounted, DWS and the Supervisory Board were made aware of the shortcomings of its practices and there was an extensive evidence trail (SEC Order, 2023). That was not the point because, as in many cases of fraud, the estimation by the organization is that it will not be caught. The SEC stated in its Order that DWS:

willfully violated Section 206(2) of the Advisers Act, which prohibits an investment adviser, directly or indirectly, from engaging “in any transaction, practice or course of business which operates as a fraud or upon any client or prospective client”. (SEC Order, 2023, ¶ 27)

As Fixler recounts, “you know this is Deutsche Bank, my eyes were open I understood they had some cultural issues...loose compliance and financial controls” (2023). The institutional logics literature suggests that organizations will move towards stabilizing by either reinforcing existing beliefs and norms, hybridizing logics or transforming to a new dominant logic (Besharov & Smith, 2014b). It is difficult to know the extent to which the scandal and ensuing fines led to either hybridization or transformation. Examining the available evidence on DWS, the notion of co-existing logics put forward by McPherson and Sauder (2013) seems plausible and it signals that fraud risk cannot be ameliorated by the existence of internal controls alone. The legal sociologist, Teubner suggests that many of today’s multinational corporations span many legal systems and are in a position to set their own legal norms almost regardless of what the law may require in a particular jurisdiction (Teubner, 2017).

The content in annual reports after the scandal evidences that, at least in terms of misstatements, DWS significantly toned down its claims about the integration of ESG DWS Group, (2000, 2001, 2002, 2003, 2004, 2005). The greater regulatory focus on sustainability (ESMA, 2024) have also underscored the governance responsibilities of boards to provide greater oversight of sustainability-related claims.

It has been a concern for some time of standard setters and regulators that sustainability standards are difficult to apply and even harder to provide assurance against in terms of ‘suitable criteria’ (International Auditing and Assurance Standards Board, 2024). Similarly, the European Commission has well-documented the shortcomings of the EU Taxonomy and its ability to be applied as an effective screening of sustainability-related investments (European Commission, 2025).

At an institutional level, the DWS Scandal highlights that fraud finds opportunities within the prevailing institutional logic, particularly in organizational cultures that tolerate ‘loose’ compliance regimes. It seems no single policy prescription will eliminate fraud risk related to false or misleading claims about sustainability. As Fixler suggests, “we need to ensure there’s enforcement. Having tons of regulation without supervision and enforcement is what I call ‘form over substance’” (2023).

Conclusion

The DWS Scandal and the resulting enforcement action clearly call out ‘greenwashing’. Making untrue claims about the sustainability credentials of products is not just akin to fraud but is actually sanctioned as fraudulent. It illustrates how institutional logic within an organization can encode and thereby legitimize beliefs and practices that normalize wrongdoing. As Pierson (2004) suggests, the historical context is significant in understanding the dynamics of how organizations and their leadership characterize and construct fraud risk in their operations. As the DWS case demonstrates, there can be a considerable lag between exploiting

ambiguity in the market place and regulators setting guardrails on what constitutes acceptable practice.

The ambiguity and lack of enforcement are now giving way to greater scrutiny and consequences for evidencing the ‘greenness’ of investment products. However, this remains a contested space, and there remain ample opportunities to mislead investors. The absence of consequences for actions or misstatements further accelerates how teams and individuals in investment management companies interpret what is ‘acceptable’ (Bazerman & Tenbrunsel, 2011). The most salient point is that responsibility for strong and effective internal controls and cultural practices remains the province of those charged with governance of the corporation. It is not the role of regulators to prevent fraud, but for the directors of a company to take the necessary steps to ensure that policies and procedures are followed and that there are consequences for material breaches; otherwise, the company is exposed to even greater litigation risk, which may result in substantial loss of business value.

The DWS case presented in this paper illustrates vividly the play of competing institutional logics and ensuing actions within the organization to stabilize and reclaim the narrative. It underscores the precarious position of whistleblowers, like Fixler, who call out bad behavior, and while individuals have agency within a corporation, there are severe consequences for attempting to question the prevailing ‘script’ which is dominant in shaping behavior and actions.

This paper contributes to the literature by elucidating the changing construction of fraud and the institutional factors that can contribute to its prevalence and normalization within a given professional context. As with any case study, there are limitations both in terms of the sources of evidence available and in the generalizability of results. Future research is needed to better understand the institutional factors that drive not only breaches of the law but also the ethical implications for the accounting profession and the basis for forming and supporting professional judgments. As with any claims and assertions made by management of a corporation, responsibility for and nature of the evidence required to support them is historically bound and continues to evolve.

References

- Abela, M. (2025). Is greenwashing fraud? In D. Malinić & S. Vučković Milutinović (Eds.), *Contemporary forensic accounting trends in combating financial fraud: Global and national perspectives* (pp. 55–70). University of Belgrade, Faculty of Economics and Business
- Anthropic. (2026). Claude Sonnet 5 [Large language model]. <https://claude.ai>
- Bartz, T. (2021, September 1). Former Deutsche Bank Executive on Green Investments: “The Sustainability Propaganda Got Completely Out of Control.” *Der Spiegel*. <https://www.spiegel.de/international/business/former-deutsche-bank->

executive-on-green-investments-the-sustainability-propaganda-got-completely-out-of-control-a-a07d3744-e79f-44bc-ab12-97b70c000b86?sara_ref=re-xx-cp-sh

Bazerman, M. H., & Tenbrunsel, A. E. (2011). *Blind spots: Why we fail to do what's right and what to do about it*. In *Blind Spots*. Princeton University Press.

Berger, P. L., & Luckmann, T. (1991). *The social construction of reality: A treatise in the sociology of knowledge*. Penguin Books.

Besharov, M. L., & Smith, W. K. (2014). Multiple institutional logics in organizations: Explaining their varied nature and implications. *Academy of Management Review*, 39(3), 364–381. <https://doi.org/10.5465/amr.2011.0431>

BEUC. (2025). *Sustainable Finance, Too Green to be True?*

Brown, A. D., Colville, I., & Pye, A. (2015). Making Sense of Sensemaking in Organization Studies. *Organization Studies*, 36(2), 265–277. <https://doi.org/10.1177/0170840614559259>

Can, O., & Turker, D. (2025). Institutional pressures and greenwashing in social responsibility: reversing the link with hybridization capability. *Management Decision*, 63(1), 187–216. <https://doi.org/10.1108/MD-10-2023-1790>

Dimaggio, P. J., & Powell, W. W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, 48(2), 147–160.

Durand, R., & Thornton, P. H. (2018). Categorizing institutional logics, institutionalizing categories: A review of two literatures. *Academy of Management Annals*, 12(2), 631–658. <https://doi.org/10.5465/annals.2016.0089>

DWS Group. (2021). *DWS Annual Report 2020*

DWS Group (2022), *DWS Annual Report 2021*

DWS Group (2023), *DWS Annual Report 2022*

DWS Group (2024), *DWS Annual Report 2023*

DWS Group (2025), *DWS Annual Report 2024*

DWS Group (2026), *DWS Annual Report 2025*

DWS Group. (2023, September 25). *DWS: Statement following U.S. Securities and Exchange Commission Settlements* DWS. <https://www.dws.com/en-gb/our-profile/media/media-releases/dws-statement-following-us-securities-and-exchange-commission-settlements/>

ESMA. (2024). *Final report on greenwashing: Response to the European Commission's request for input on "greenwashing risks and the supervision of sustainable finance policies"*. https://www.esma.europa.eu/sites/default/files/2024-06/ESMA36-287652198-2699_Final_Report_on_Greenwashing.pdf

European Commission. (2019). *The European Green Deal*. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en

European Commission. (2025). *EU Taxonomy – Review of Climate and Environmental Delegated Acts*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=intcom%3AAres%282025%299611285>

European Parliament. (2020). Briefing - Implementation Appraisal Non-financial Reporting Directive. [https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/654184/EPRS_BRI\(2020\)654184_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2020/654184/EPRS_BRI(2020)654184_EN.pdf)

Fixler, D. (2023, April 20). Conferences 2022/23 | Sustainability 20|30. Keynote Day 2. <https://sustentabilidade.negocios.pt/conferencias-2022-23/>

Fixler, D., & Eccles, R. (2025). Greenwashing: In Conversation with a Whistleblower. *Accountability for a Sustainable World Quarterly*, 133–139. <https://doi.org/10.7274/28583123.V1>

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, 12(2), 219–245. <https://doi.org/10.1177/1077800405284363>

Friedland, R., & Alford, R. R. (2013). Bringing society back in: Symbols, practices, and institutional contradictions. In W. W. Powell & P. J. DiMaggio (Eds.), *The new institutionalism in organizational analysis* (pp. 232–263). University of Chicago.

Hartmann, B., Marton, J., & Söderström, R. (2018). The Improbability of Fraud in Accounting for Derivatives: A Case Study on the Boundaries of Financial Reporting Compliance. *European Accounting Review*, 27(5), 845–873. <https://doi.org/10.1080/09638180.2018.1494022>

Hellmann, A., Perera, H., & Patel, C. (2013). Continental European accounting model and accounting modernization in Germany. *Advances in Accounting*, 29(1), 124–133. <https://doi.org/10.1016/j.adiaac.2013.02.002>

International Auditing and Assurance Standards Board. (2024). *ISSA 5000 Sustainability Assurance-Basis for Conclusions*. <https://ifacweb.blob.core.windows.net/publicfiles/2025-01/IAASB-ISSA-5000-Sustainability-Assurance-Basis-for-Conclusions.pdf>

Jastram, S. M., & Foersterling, J. (2024). Clarifying ends in institutional decoupling theory: insights from social impact research. *Journal of Business Economics*, 94(3), 501–523. <https://doi.org/10.1007/s11573-023-01170-8>

Jepperson, R. L., & Meyer, J. W. (2021). *Institutional Theory*. Cambridge University Press.

- Kauppi, K., & Hannibal, C. (2017). Institutional pressures and sustainability assessment in supply chains. *Supply Chain Management*, 22(5), 458–472. <https://doi.org/10.1108/SCM-01-2017-0004>
- Kowsmann, P., & Brown, K. (2021, August 1). Fired Executive Says Deutsche Bank's DWS Overstated Sustainable-Investing Efforts. *Wall Street Journal*.
- Krippendorff, Klaus. (2013). *Content analysis: an introduction to its methodology* (3rd ed.). SAGE.
- Kuhn, T. S. (1987). What are scientific revolutions? In L. Kruger, L. J. Daston, & M. Heidelberger (Eds.), *The probabilistic revolution, Vol. 1: Ideas in history* (pp. 7–22). MIT Press.
- Kurpierz, J. R., & Smith, K. (2020). The greenwashing triangle: adapting tools from fraud to improve CSR reporting. *Sustainability Accounting, Management and Policy Journal*, 11(6), 1075–1093. <https://doi.org/10.1108/SAMPJ-10-2018-0272>
- Marquis, C., Toffel, M. W., & Zhou, Y. (2016). Scrutiny, norms, and selective disclosure: A global study of greenwashing. *Organization Science*, 27(2), 483–504. <https://doi.org/10.1287/orsc.2015.1039>
- Matussek, K., & Arons, S. (2022, January 24). Ex-DWS manager Fixler loses labor suit over greenwashing claim. *Bloomberg*. <https://news.bloomberglaw.com/banking-law/ex-dws-manager-fixler-loses-labor-suit-over-greenwashing-claim>
- McPherson, C. M., & Sauder, M. (2013). Logics in Action: Managing Institutional Complexity in a Drug Court. *Administrative Science Quarterly*, 58(2), 165–196. <https://doi.org/10.1177/0001839213486447>
- Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363. <https://doi.org/10.1086/226550>
- Müller, F., & Storbeck, O. (2025, April 3). Deutsche asset manager fined €25mn over greenwashing. *Financial Times*. <https://login.avoserv2.library.fordham.edu/login?url=https://www.proquest.com/newspap>
- Mutch, A. (2019). *Reframing Institutional Logics: Substance, Practice and History*. Routledge.
- Palma, S., & Storbeck, O. (2023, September 26). Deutsche Bank's DWS agrees \$19mn settlement after “greenwashing” probe. *Financial Times*. <https://www.proquest.com/newspapers/deutsche-banks-dws->
- Perez, O., & Verbruggen, P. (2026). Green defaults in sustainable finance: Compliance challenges at the intersection of private and public law. *Journal of Environmental Law*, 38(2), 171–195. <https://doi.org/10.1093/jel/eqag003>

Pierson, Paul. (2004). *Politics in Time : History, Institutions, and Social Analysis*. Princeton University Press.

SEC Order (September 25, 2023).

Securities Exchange Commission. (2025, September 23). Deutsche Bank Subsidiary DWS to Pay \$25 Million for Anti-Money Laundering Violations and Misstatements Regarding ESG Investments. <https://www.sec.gov/newsroom/press-releases/2023-194>

Shy, O. (2026). Whistleblowers and financial fraud. *Journal of Financial Stability*, 82, 101494. <https://doi.org/10.1016/j.jfs.2025.101494>

Teubner, G. (2017). Corporate Codes in the Varieties of Capitalism: How Their Enforcement Depends on the Differences Among Production Regimes. *Indiana Journal of Global Legal*, (24), 81–97. <https://muse.jhu.edu/article/658268>

Thornton, P. H., & Ocasio, W. (2008). Institutional Logics. In R. Greenwood, C. Oliver, R. Suddaby, K. Sahlin, Thornton P.H., & W. Ocasio (Eds.), *The SAGE Handbook of Organizational Institutionalism* (pp. 99–128). SAGE Publications Ltd. <https://doi.org/10.4135/9781849200387.n4>

Thornton, P. H., Ocasio, William., & Lounsbury, Michael. (2012). *The Institutional Logics Perspective : A New Approach to Culture, Structure, and Process*, Oxford University Press.

Tight, M. (2010). The curious case of case study: A viewpoint. *International Journal of Social Research Methodology*, 13(4), 329–339. <https://doi.org/10.1080/13645570903187181>

Tindera, M., & Temple-West, P. (2022, June 8). Blowing the whistle on ESG. *FT Behind the Money*. <https://www.ft.com/content/c1fd5ed0-e3f1-4c54-91f7-f0ad5a10af9a>

Weick, K. E. (1995). *Sensemaking in Organizations*. Sage Publications.

Yin, R. K. (1994). Discovering the future of the case study method in evaluation research. *Evaluation Practice*, 15(3), 283–290. [https://doi.org/10.1016/0886-1633\(94\)90023-X](https://doi.org/10.1016/0886-1633(94)90023-X)

Yu, J., & Hwang, Y. S. (2026). Behind the green veil: board characteristics, greenwashing, and the fraud triangle. *Humanities and Social Sciences Communications*, 13(1), 632. <https://doi.org/10.1057/s41599-026-06977-8>

Bobana Čegar Babić
University of Banja Luka, Faculty of
Economics
ORCID: 0009-0008-2844-6895

Tajana Serdar Raković
University of Banja Luka, Faculty of
Economics
ORCID: 0000-0001-9177-608X

FINANCIAL STATEMENT MANIPULATION AND CAPITAL MARKET INTEGRITY: THE ROLE OF FORENSIC ACCOUNTING IN DETECTING EARNINGS MANAGEMENT

UDK 657.6:336.76

DOI <https://doi.org/10.54318/czf.2026.bcb.015>

Abstract

Capital market integrity refers to a transparent, efficient, and manipulation-free market. In this context, it is essential that investors are able to make decisions based on reliable information and high degree of confidence in financial reporting and regulatory institutions. Unlike traditional financial auditing, which provides reasonable assurance that financial statements present a fair and objective view of a company's actual financial position and performance, forensic accounting is focused on detecting potential fraud and manipulation in financial reporting, such as overstating earnings, understating liabilities, or overstating assets, as well as collecting evidence that may be used in judicial and other regulatory proceedings. The aim of this paper is to apply the Beneish M-Score, as an analytical forensic accounting model, to identify companies whose indicators point to possible earnings manipulation, and to examine the implications for investors, regulators, auditors, and other stakeholders. The analysis is based on disclosed financial information of companies whose securities are traded on stock exchanges in selected Southeast European countries: Bosnia and Herzegovina, Croatia, Serbia, Montenegro and Slovenia. Based on the obtained results, statistical methods are used to examine whether there are significant differences among the observed countries and whether there is a relationship between the Beneish M-Score and the type of issued audit opinion.

Key words: financial statements, capital market, forensic accounting, Beneish M-Score, earning management

JEL Classification: M41, G14, M42

Introduction

Reliable financial reporting supports capital-market confidence by reducing information asymmetry, improving investment decisions, and facilitating the efficient allocation of resources (Healy & Palepu, 2001; Roychowdhury, Shroff, & Verdi, 2019). Its integrity is threatened when managerial judgement is used opportunistically to alter reported results, mislead stakeholders, or influence outcomes based on accounting information (Healy & Wahlen, 1999). Earnings management may involve accrual-based choices or real activities manipulation and may progress towards financial statement manipulation and fraud. These practices can reduce earnings informativeness, distort investment decisions, and weaken confidence in financial reporting (Habib et al., 2022).

External auditing provides reasonable assurance that financial statements are free from material misstatement (IAASB, 2022). An unmodified opinion therefore does not guarantee the absence of earnings management. Forensic accounting complements auditing by identifying unusual financial patterns. The Beneish M-Score uses financial statement variables to flag potential manipulation (Beneish, 1999; Beneish, Lee, & Nichols, 2013; Maniatis, 2022), but provides a risk signal rather than proof.

Although the Beneish M-Score has been applied in individual European markets (Fenyves, Pisula, & Tarnóczy, 2023; Maniatis, 2022), comparative evidence from Southeast Europe remains limited. The relationship between model-based manipulation signals and independent audit opinions also remains under-explored. This paper applies the Beneish M-Score to publicly available data from eligible listed companies in selected Southeast European countries, addressing two research questions:

RQ1: Do Beneish M-Scores vary significantly among listed companies across the observed Southeast European countries?

RQ2: Is there a relationship between the Beneish M-Score and the type of audit opinion a company receives?

The paper makes three contributions. Empirically, it provides comparative evidence from under-researched Southeast European capital markets. Methodologically, it combines a forensic screening model with independent audit-report information. Practically, the study shows how investors, auditors, regulators, and stock exchanges can combine public accounting and audit information to identify reporting-risk signals and guide further investigation.

1. Theoretical background and previous research

Although earnings management is often discussed alongside manipulation and fraud, these concepts are not synonymous. Earnings management may occur through choices permitted by accounting standards, whereas fraudulent reporting involves

intentional material misstatements or omissions that violate reporting requirements (Dechow & Skinner, 2000). Financial statement manipulation represents a more aggressive distortion and may cross into fraud. The more difficult distinction is between legitimate accounting judgement and earnings management, as it depends largely on managerial intent (Healy & Wahlen, 1999).

1.1. Earnings Management, Financial Statement Manipulation, and Capital-Market Integrity

Earnings management may take the form of accrual-based accounting choices or real activities manipulation. Accrual-based earnings management involves changes in estimates, provisions, depreciation methods, revenue recognition, or expense timing. These choices change reported earnings without immediately altering the company's operations. Real activities manipulation affects actual business decisions. It may include offering temporary price discounts to increase sales, overproducing to reduce reported cost of goods sold, or cutting discretionary expenditure to improve short-term earnings (Roychowdhury, 2006). Managers may use either approach to meet analysts' forecasts, avoid losses, comply with debt covenants, or increase performance-related compensation (Healy & Wahlen, 1999). Survey evidence indicates that managers may sacrifice long-term economic value to meet short-term earnings targets (Graham, Harvey & Rajgopal, 2005). Managers may trade off accrual-based and real earnings management according to their relative costs, timing, and monitoring constraints (Zang, 2012). The extent of real earnings management is also shaped by governance, capital-market, and regulatory factors (Habib et al., 2022).

The prevalence and consequences of earnings management vary across institutional settings. International evidence shows that earnings management is less prevalent in countries with stronger investor protection and legal enforcement (Leuz, Nanda & Wysocki, 2003). Evidence from European firms further indicates that earnings management is lower among public than private firms and in countries with stronger legal enforcement (Burgstahler, Hail & Leuz, 2006). US enforcement cases show that financial misrepresentation can result in legal penalties and substantial market-value losses. Reputational damage accounts for a large share of the market penalty (Karpoff, Lee & Martin, 2008).

Nevertheless, accounting discretion is not necessarily opportunistic. Discretionary accruals may contain information about future firm performance (Subramanyam, 1996). Changes in the stock prices of firms with greater income smoothing contain more information about future earnings than corresponding changes for firms with less smoothing (Tucker & Zarowin, 2006). These findings suggest that accounting discretion cannot be classified as harmful solely from observed financial patterns. Analytical indicators must therefore be interpreted considering company circumstances and the wider institutional environment.

1.2. Forensic Accounting, the Beneish M-Score and Audit Opinions

Statutory auditing and forensic accounting differ in scope and purpose. A statutory audit addresses the financial statements within the framework of materiality and reasonable assurance. Forensic accounting is problem-specific and examines particular anomalies, allegations or suspected schemes. It may also collect evidence for regulatory or legal proceedings (IAASB, 2022; Singleton & Singleton, 2010).

Beneish (1999) developed the M-Score using data from US public companies that had overstated earnings and a control group of non-manipulators. Its indicators capture possible reporting distortions and conditions associated with incentives to manipulate earnings.

European applications produce markedly different results. Maniatis (2022) classified 17.5% of the examined Greek companies as potential manipulators, while Fenyves, Pisula, and Tarnóczy (2023) reported proportions of approximately 46-52% in Hungary. Svabova et al. (2020) developed a locally adapted model for Slovak companies. Comporek (2020) found that the original Beneish model identified only 40.74% of 27 Polish companies sanctioned for reporting irregularities. Although these studies differ in their samples and methods, their findings indicate that model performance may vary across markets.

Evidence concerning auditor reporting is also mixed. Francis and Krishnan (1999) found that high-accrual firms were more likely to receive modified reports for asset-realisation uncertainties and going-concern problems. Bartov, Gul, and Tsui (2000) found that the absolute magnitude of discretionary accruals estimated by the cross-sectional Jones models was positively associated with the likelihood of receiving a qualified audit report, even after controlling for litigation risk and earnings performance. Fakhfakh and Jarboui (2021) similarly found that earnings management increased the likelihood of a modified opinion among Tunisian listed firms. Conversely, Butler, Leone and Willenborg (2004) found that the association between abnormal accruals and modified reports was concentrated among companies with going-concern opinions and largely reflected financial distress. Tsipouridou and Spathis (2014) found no relationship between audit opinions and earnings management in Greece. These conflicting findings leave the relationship between manipulation risk and audit opinions empirically unresolved.

2. Research Methodology

This study uses a quantitative, comparative, and cross-sectional approach, leveraging secondary financial data to explore its research questions.

2.1 Research Design and Hypotheses

The study focuses on companies listed in select Southeast European six capital markets: Bosnia and Herzegovina (Republic of Srpska and Federation of BiH),

Serbia, Croatia, Montenegro and Slovenia. The Beneish M-Score is calculated using financial data for 2025 and corresponding comparative data for 2024.

Based on previous evidence of cross-country differences in manipulation risk (Hasan et al., 2017) and the relationship between earnings management and audit opinions (Fakhfakh & Jarboui, 2021), the following hypotheses are tested:

H1: The distribution of Beneish M-Scores varies significantly among listed companies across the observed Southeast European countries.

H2: Companies with modified audit opinions tend to have higher Beneish M-Scores compared to companies with unmodified audit opinions.

2.2 Sample and Data Sources

The initial sample comprised companies listed on the Banja Luka, Sarajevo, Belgrade, Zagreb, Podgorica, and Ljubljana stock exchanges and included in their respective main stock market indices. The analysis used publicly available financial statements for 2024 and 2025, and independent auditors' reports.

Companies were included if the financial data required to calculate all eight Beneish M-Score components were available and comparable across both years. Banks, insurance companies, investment funds, and other financial institutions were excluded because their financial statements and operating characteristics differ from those of the non-financial companies for which the original model was developed (Beneish, 1999). Companies were also excluded if they had incomplete data, inconsistent reporting periods, or undefined ratios caused by missing or zero denominators.

The final M-Score sample comprised 97 companies: Bosnia and Herzegovina – 32 (Banjaluka Stock Exchange -11 and Sarajevo Stock Exchange - 21), Serbia - 19, Croatia - 22, Montenegro - 17, and Slovenia - 7. The audit-opinion analysis included only companies with publicly available independent audit reports. An unavailable audit report was treated as missing information, not as a separate opinion category. However, an additional analysis was conducted to assess whether the exclusion of these companies from the sample could affect the final decision regarding the acceptance or rejection of the hypotheses. Although, the sample constructed in this manner is unbiased and includes companies with the most liquid securities, the representation of individual countries in the overall sample varies considerably, solely reflecting how stock market indices are constructed in each of the capital markets examined.

2.3 Variable Measurement and Statistical Analysis

The central focus of this study is the Beneish M-Score, which serves as a probabilistic tool to assess the likelihood of displaying financial characteristics associated with earnings manipulation. Developed by Beneish (1999), this model incorporates eight variables:

$$M - Score = -4.840 + 0.920 \times DSRI + 0.528 \times GMI + 0.404 \times AQI + 0.892 \times SGI \\ + 0.115 \times DEPI - 0.172 \times SGAI + 4.679 \times TATA - 0.327 \times LVGI$$

Table 1 presents the calculation and interpretation of each indicator. In all formulas, t denotes 2025 and $t-1$ denotes 2024.

Table 1: Components of the Beneish M-Score

Indicator	Linear formula	Interpretation
DSRI - Days' Sales in Receivables Index	$\frac{Receivables_t / Sales_t}{Receivables_{t-1} / Sales_{t-1}}$	Identifies whether receivables increased disproportionately relative to sales, potentially indicating premature or inflated revenue recognition.
GMI – Gross Margin Index	$\frac{Gross\ profit_{t-1} / Sales_{t-1}}{Gross\ profit_t / Sales_t}$	Measures deterioration in gross margin, which may increase pressure on management to manipulate earnings.
AQI - Asset Quality Index	$1 - \left(\frac{Current\ assets_t + PPE_t}{Total\ assets_t} \right) \\ 1 - \left(\frac{Current\ assets_{t-1} + PPE_{t-1}}{Total\ assets_{t-1}} \right)$	Measures changes in the proportion of assets whose future benefits are less certain and may indicate increased cost capitalisation
SGI -Sales Growth Index	$\frac{Sales_t}{Sales_{t-1}}$	Captures sales growth, which may create pressure to maintain performance and market expectations.
DEPI - Depreciation Index	$\frac{Dep_{t-1} / (Dep_{t-1} + PPE_{t-1})}{Dep_t / (Dep_t + PPE_t)}$	Identifies a possible reduction in the depreciation rate that may increase reported earnings
SGAI -Sales, General and Administrative Expenses Index	$\frac{SG\&A\ Expenses_t / Sales_t}{SG\&A\ Expenses_{t-1} / Sales_{t-1}}$	Measures changes in administrative and selling expenses relative to sales and may signal deteriorating operating efficiency.
LVGI- Leverage Index	$\frac{Total\ liabilities_t / Total\ assets_t}{Total\ liabilities_{t-1} / Total\ assets_{t-1}}$	Measures changes in leverage and the potential pressure arising from debt obligations and contractual constraints.
TATA -Total Accruals to Total Assets	$\frac{Profit\ before\ tax_t - CFO_t}{Total\ assets_t}$	Measures the extent to which reported earnings are supported by operating cash flows; higher accruals may indicate lower earnings quality.

Source: Beneish (1999).

Notes: Given differences between US-GAAP and IAS/IFRS reporting and the inconsistent availability of income from continuing operations, TATA was calculated using profit before tax, while the original model coefficients were retained.

CA - current assets; PPE - property, plant and equipment; TA - total assets; COGS - cost of goods sold; Dep - depreciation expense; SG&A - selling, general and administrative expenses; CL - current liabilities; LTD - long-term debt; CFO - net cash flow from operating activities.

The continuous M-Score is used as the principal metric. Binary classifications based on the -2.22 and -1.78 thresholds are used as sensitivity checks (Beneish, 1999; Beneish et al., 2013). A score below -2.22 (i.e. more negative than -2.22) indicates that company most probably is a non-manipulator, between -2.22 and -1.78 is grey area, which means that the company might be manipulator, and the score above -1.78 suggests that the company most likely does earnings manipulation. However, it should be kept in mind that M-score provides a signal or indicates a potential risk but does not constitute proof that a company has actually manipulated its earnings.

Audit opinion is coded as a binary variable: unmodified opinions as 0 and modified opinions as 1. Modified opinions include qualified and adverse opinions and disclaimers of opinion. An emphasis-of-matter paragraph accompanying an otherwise unmodified opinion does not change its classification. Unavailable audit reports are treated as missing data.

Descriptive statistics were calculated for the total sample and by market. As M-Score distributions were non-normal, H1 was tested using the Kruskal–Wallis test and H2 using the Mann–Whitney U test. The latter also assessed differences between companies with and without available audit reports. All tests were conducted in SPSS at the 5% significance level.

3. Results

The empirical analysis begins with descriptive statistics for the Beneish M-Score, reported for the total sample and separately by country. The distributions are also examined for normality, extreme values, and potentially influential observations.

Table 2: Descriptive Analysis of the Beneish M-Score by Country

Country	N	Min	Max	Median	IQR	M-score (% of companies)		
						< -2.22	-2.22 to -1.78	> -1.78
BiH – RS	11	-3.1	-1.11	-2.34	0.88	63.6	27.3	9.1
BiH – FBiH	21	-5.13	7.71	-2.38	1.19	66.7	14.3	19.0
Serbia	19	-3.3	5.9	-2.53	0.97	63.2	26.3	10.5
Croatia	22	-4.02	-0.29	-2.58	0.79	68.2	22.7	9.1
Montenegro	17	-8.1	0.33	-2.62	0.74	82.4	5.9	11.8
Slovenia	7	-3.09	-2.2	-2.46	0.44	85.7	14.3	0.0

Source: Authors' analysis

To test H1, the following statistical null hypothesis was specified:

H1₀: Beneish M-Score distributions do not differ across the observed countries.

H1₁: Beneish M-Score distributions differ significantly between at least two of the observed countries.

As descriptive statistics (presented in Table 2) indicate, there is relatively similar median Beneish M-Scores across the observed countries, ranging from -2.62 in Montenegro to -2.34 in BiH-RS. However, some differences in dispersion can be observed, with the highest IQR recorded for BiH-FBiH (1.19) and the lowest for Slovenia (0.44). The presence of extreme values is particularly evident in BiH-FBiH and Serbia, as reflected by their wide minimum–maximum ranges.

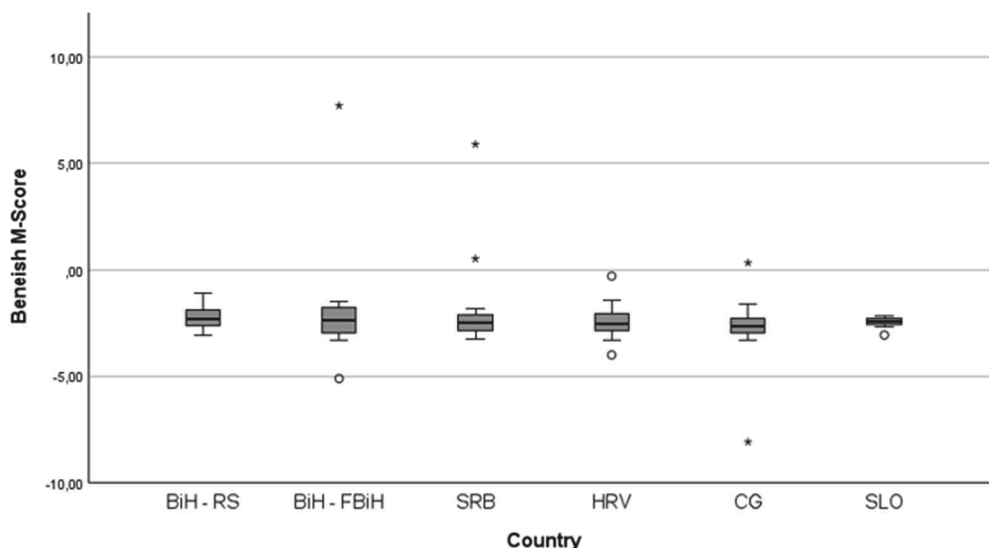
Given that the M-Score does not follow a normal distribution, the hypotheses were tested using the Kruskal–Wallis test (reported in Table 3).

Table 3. Independent-Samples Kruskal-Wallis Test, Summary

Total N	97
Test Statistic	1.874 ^a
Degree of Freedom	5
Asymptotic Sig. (2-sided test)	.866
a. The test statistic is adjusted for ties.	

Source: Authors' analysis

The Kruskal–Wallis test did not reveal statistically significant differences in Beneish M-Scores across the six observed Southeast European markets ($H = 1.874$, $df = 5$, $p = 0.866$). Therefore, H1₀ cannot be rejected and H1 is not supported, indicating that the distribution of Beneish M-Scores does not differ significantly across the countries included in the sample. This finding is also consistent with the boxplot (Chart 1), which shows relatively similar median M-Scores and interquartile ranges across the observed countries, despite the presence of several extreme observations.

Chart 1. Independent-Samples Kruskal-Wallis Test

Source: Authors' analysis

Regarding H2, the analysis investigates whether companies that did not receive an unmodified audit opinion exhibit higher M-Score values, which would indicate a greater likelihood of financial statement manipulation. To test H2, the following statistical null hypothesis was specified:

H2₀: There is no statistically significant difference in Beneish M-Scores between companies with modified and unmodified audit opinions.

H2₁: Companies with modified audit opinions have significantly higher Beneish M-Scores than companies with unmodified audit opinions.

Since the M-Score does not follow a normal distribution, is measured as a continuous variable, and the comparison involves two independent groups (0 = unmodified opinion; 1 = modified opinion), the Mann–Whitney U test is used. As previously noted, the testing of this hypothesis currently excludes 21 companies for which independent auditors' reports were not publicly available. Table 4 presents descriptive statistics.

Table 4. Descriptive Analysis of the Beneish M-Score by Auditor's opinion

Opinion	N	Min	Max	Median	IQR
Unmodified	64	-8.10	5.90	-2.59	0.70
Modified	12	-3.10	-1.11	-2.44	0.51

Source: Authors' analysis

Companies with modified audit opinions exhibit a slightly higher median M-Score than companies with unmodified audit opinions. However, the observed difference is not statistically significant, as confirmed by the Mann–Whitney U test results reported in Table 5.

Table 5. Independent-Samples Mann-Whitney U Test Summary

Total N	76
Mann-Whitney U	421.000
Wilcoxon W	499.000
Test Statistic	421.000
Standard Error	70.200
Standardized Test Statistic	.527
Asymptotic Sig. (2-sided test)	.598

Source: Authors' analysis

Accordingly, at the 5 percent significance level, the results indicate that there is no statistically significant difference in Beneish M-Scores between companies receiving modified audit opinions and those receiving unmodified audit opinions. Although companies with modified audit opinions exhibit a higher mean rank of Beneish M-Scores than companies with unmodified audit opinions (41.58 vs. 37.92), the difference is not statistically significant (Mann–Whitney U = 421.000, $p = 0.598$). Therefore, H_{2a} cannot be rejected, and H_2 is not supported.

Since 21 companies were excluded from the analysis due to the unavailability of published independent audit reports, representing approximately 22 percent of the sample, the following section examines whether there are meaningful differences in Beneish M-Score values between these companies and those for which audit reports were available, as indicated by the descriptive statistics presented in Table 6.

Table 6. Descriptive Analysis of the Beneish M-Score by availability of auditor's opinion

Opinion	N	Min	Max	Median	IQR
Available	76	-8.10	5.90	-2.55	0.66
Not available	21	-5.13	7.71	-2.34	1.30

Source: Authors' analysis

Companies without available audit opinions exhibit a somewhat higher median M-Score than companies with available audit opinions. Moreover, the interquartile range is substantially larger for this group, suggesting greater variability in M-Score

values. Nevertheless, the Mann–Whitney U test indicates that the observed differences are not statistically significant, as shown in Table 7.

Table 7. Independent-Samples Mann-Whitney U Test Summary

Total N	97
Mann-Whitney U	895.000
Wilcoxon W	1126.000
Test Statistic	895.000
Standard Error	114.167
Standardized Test Statistic	.850
Asymptotic Sig. (2-sided test)	.396

Source: Authors' analysis

Although companies with unavailable audit opinions had a somewhat higher mean rank (53.62) than companies with available audit opinions (47.72), the difference was not statistically significant. This suggests that the exclusion of observations with unavailable audit opinions is unlikely to have introduced a systematic difference in Beneish M-Scores between the two groups.

Conclusion

In the observed Southeast European countries, 70 percent of companies have Beneish M-Score values below -2.22 . According to this indicator, these companies are associated with a low likelihood of earnings management and financial statement manipulation. Although some variations are observed across countries, these differences are not statistically significant. Identifying the factors underlying such results opens an avenue for further research. It may be assumed that these findings partly reflect the similar historical and institutional development of the observed countries, as well as comparable regulatory frameworks. In addition, their financial systems share a similar structure, being predominantly bank-oriented. Compared with some developed countries in Western and Central Europe, capital markets in Southeast Europe are less liquid, and investor pressure is therefore less pronounced.

The findings further indicate that companies receiving modified audit opinions exhibit higher M-Score values and mean ranks than those receiving unmodified audit opinions. Nevertheless, the observed difference is not statistically significant. Since unmodified opinions dominate the sample (84 percent), and the proportion of companies audited by Big Four firms is relatively limited (less than 30 percent), further analysis incorporating auditor characteristics may provide a more nuanced assessment of the relationship between audit quality and the likelihood of earnings manipulation.

The absence of a statistically significant relationship between audit opinions and Beneish M-Scores suggests that traditional audit opinions, when considered alone, may have limited capacity to identify potential financial statement manipulation. This underscores the complementary role of forensic accounting, which can provide more targeted analytical and investigative procedures for companies identified by screening models such as the Beneish M-Score as potentially exhibiting manipulative reporting behaviour.

References

- Bartov, E., Gul, F. A., & Tsui, J. S. L. (2000). Discretionary-accruals models and audit qualifications. *Journal of Accounting and Economics*, 30(3), 421–452. [https://doi.org/10.1016/S0165-4101\(01\)00015-5](https://doi.org/10.1016/S0165-4101(01)00015-5)
- Beneish, M. D. (1999). The detection of earnings manipulation. *Financial Analysts Journal*, 55(5), 24–36. <https://doi.org/10.2469/faj.v55.n5.2296>
- Beneish, M. D., Lee, C. M. C., & Nichols, D. C. (2013). Earnings manipulation and expected returns. *Financial Analysts Journal*, 69(2), 57–82. <https://doi.org/10.2469/faj.v69.n2.1>
- Butler, M., Leone, A. J., & Willenborg, M. (2004). An empirical analysis of auditor reporting and its association with abnormal accruals. *Journal of Accounting and Economics*, 37(2), 139–165. [https://doi.org/10.1016/S0165-4101\(03\)00061-2](https://doi.org/10.1016/S0165-4101(03)00061-2)
- Burgstahler, D. C., Hail, L., & Leuz, C. (2006). The importance of reporting incentives: Earnings management in European private and public firms. *The Accounting Review*, 81(5), 983–1016. <https://doi.org/10.2308/accr.2006.81.5.983>
- Comporek, M. (2020). The effectiveness of the Beneish model in the detection of accounting violations: The example of companies sanctioned by the Polish Financial Supervision Authority. *Research Papers of Wroclaw University of Economics and Business*, 64(10), 18–30. <https://doi.org/10.15611/pn.2020.10.02>
- Dechow, P. M., & Skinner, D. J. (2000). Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons*, 14(2), 235–250. <https://doi.org/10.2308/acch.2000.14.2.235>
- Fakhfakh, I., & Jarboui, A. (2021). The moderating role of audit quality on the relationship between auditor reporting and earnings management: Empirical evidence from Tunisia. *EuroMed Journal of Business*, 16(4), 416–430. <https://doi.org/10.1108/EMJB-03-2020-0024>
- Fenyves, V., Pisula, T., & Tarnóczy, T. (2023). Investigation of accounting manipulation using the Beneish model: Hungarian case. *Economics and Sociology*, 16(4), 347–363. <https://doi.org/10.14254/2071-789X.2023/16-4/18>

- Francis, J. R., & Krishnan, J. (1999). Accounting accruals and auditor reporting conservatism. *Contemporary Accounting Research*, 16(1), 135–165. <https://doi.org/10.1111/j.1911-3846.1999.tb00577.x>
- Graham, J. R., Harvey, C. R., & Rajgopal, S. (2005). The economic implications of corporate financial reporting. *Journal of Accounting and Economics*, 40(1–3), 3–73. <https://doi.org/10.1016/j.jacceco.2005.01.002>
- Habib, A., Ranasinghe, D., Wu, J. Y., Biswas, P. K., & Ahmad, F. (2022). Real earnings management: A review of the international literature. *Accounting & Finance*, 62(4), 4279–4344. <https://doi.org/10.1111/acfi.12968>
- Hasan, M. S., Omar, N., Barnes, P., & Handley-Schachler, M. (2017). A cross-country study on manipulations in financial statements of listed companies: Evidence from Asia. *Journal of Financial Crime*, 24(4), 656–677. <https://doi.org/10.1108/JFC-07-2016-0047>
- Healy, P. M., & Palepu, K. G. (2001). Information asymmetry, corporate disclosure, and the capital markets: A review of the empirical disclosure literature. *Journal of Accounting and Economics*, 31(1–3), 405–440. [https://doi.org/10.1016/S0165-4101\(01\)00018-0](https://doi.org/10.1016/S0165-4101(01)00018-0)
- Healy, P. M., & Wahlen, J. M. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting Horizons*, 13(4), 365–383. <https://doi.org/10.2308/acch.1999.13.4.365>
- International Auditing and Assurance Standards Board. (2022). *International Standard on Auditing 200: Overall objectives of the independent auditor and the conduct of an audit in accordance with International Standards on Auditing*. In *Handbook of international quality management, auditing, review, other assurance, and related services pronouncements* (Vol. 1). International Federation of Accountants.
- Karpoff, J. M., Lee, D. S., & Martin, G. S. (2008). The cost to firms of cooking the books. *Journal of Financial and Quantitative Analysis*, 43(3), 581–611. <https://doi.org/10.1017/S0022109000004221>
- Leuz, C., Nanda, D., & Wysocki, P. D. (2003). Earnings management and investor protection: An international comparison. *Journal of Financial Economics*, 69(3), 505–527. [https://doi.org/10.1016/S0304-405X\(03\)00121-1](https://doi.org/10.1016/S0304-405X(03)00121-1)
- Maniatis, A. S. (2022). Detecting the probability of financial fraud due to earnings manipulation in companies listed in Athens Stock Exchange Market. *Journal of Financial Crime*, 29(2), 603–619. <https://doi.org/10.1108/JFC-04-2021-0083>

- Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of Accounting and Economics*, 42(3), 335–370. <https://doi.org/10.1016/j.jacceco.2006.01.002>
- Roychowdhury, S., Shroff, N., & Verdi, R. S. (2019). The effects of financial reporting and disclosure on corporate investment: A review. *Journal of Accounting and Economics*, 68(2–3), Article 101246. <https://doi.org/10.1016/j.jacceco.2019.101246>
- Singleton, T. W., & Singleton, A. J. (2010). *Fraud auditing and forensic accounting* (4th ed.). Wiley.
- Subramanyam, K. R. (1996). The pricing of discretionary accruals. *Journal of Accounting and Economics*, 22(1–3), 249–281. [https://doi.org/10.1016/S0165-4101\(96\)00434-X](https://doi.org/10.1016/S0165-4101(96)00434-X)
- Svabova, L., Kramarova, K., Chutka, J., & Strakova, L. (2020). Detecting earnings manipulation and fraudulent financial reporting in Slovakia. *Oeconomia Copernicana*, 11(3), 485–508. <https://doi.org/10.24136/oc.2020.020>
- Tsipouridou, M., & Spathis, C. (2014). Audit opinion and earnings management: Evidence from Greece. *Accounting Forum*, 38(1), 38–54. <https://doi.org/10.1016/j.accfor.2013.09.002>
- Tucker, J. W., & Zarowin, P. A. (2006). Does income smoothing improve earnings informativeness? *The Accounting Review*, 81(1), 251–270. <https://doi.org/10.2308/accr.2006.81.1.251>
- Zang, A. Y. (2012). Evidence on the trade-off between real activities manipulation and accrual-based earnings management. *The Accounting Review*, 87(2), 675–703. <https://doi.org/10.2308/accr-10196>

Ivana Vuković
Faculty of Economics University of
Kragujevac
ORCID: 0009-0007-7595-0624

Dejan Jovanović
Faculty of Economics University of
Kragujevac
ORCID: 0000-0002-0424-0656

Mirjana Todorović
Faculty of Economics University of
Kragujevac
ORCID: 0000-0003-0209-6694

THE RELATIONSHIP BETWEEN THE ENVIRONMENTAL COMPONENT OF ESG REPORTING AND THE FINANCIAL PERFORMANCE OF EUROPEAN COMPANIES

UDK 502.131.1:657.3

DOI <https://doi.org/10.54318/czf.2026.iv.016>

Abstract

Changes in the business environment and growing stakeholder expectations are directing companies' attention toward the responsible management of the impacts of their operations on the environment, society, and corporate governance. These aspects of business operations are integrated through the ESG concept, while ESG reporting represents an important mechanism for the transparent presentation of the aforementioned aspects of business operations and a basis for assessing their relationship with companies' financial performance. Based on the foregoing, the main objective of this paper is to determine whether there is a relationship between the E component of ESG reporting, expressed through the environmental component index calculated as the ratio of the total number of disclosed environmental indicators to the total number of environmental reporting indicators defined in accordance with the GRI Standards, and companies' financial performance, expressed by the return on assets and return on equity. The research methodology is based, on the one hand, on a qualitative approach, which was applied in the analysis of relevant literature and the formulation of research hypotheses, and, on the other hand, on statistical analysis using descriptive statistics and Spearman's correlation coefficient. The research included companies headquartered in fifteen European countries that received the highest rating (an A score) within the publicly available database of the Carbon Disclosure Project (CDP). The sample comprised 100 large companies over a five-year period from 2019 to 2023, out of a total of 227 European companies listed in the CDP database that were ranked highest according to CDP criteria and received an A score. The results of the research indicated that there is no statistically significant relationship between the E component of ESG reporting and the financial performance of the observed European companies. These results indicate that the mere existence of ESG reporting is not a sufficient condition for the

existence of a relationship and the achievement of an impact on companies' financial performance, but that the quality of reports, transparency in the presentation of information, and companies' genuine commitment to implementing sustainable practices play a key role. In addition, the results also indicate the importance of distinguishing transparent reporting from greenwashing practices, which may create an impression of greater commitment to sustainability than is actually reflected in companies' operations.

Key words: sustainability, ESG, environmental component, financial performance, greenwashing

JEL Classification: M41, Q01, Q56

Introduction

The global challenges faced by the modern world have led to changes in approaches to business and an increasing focus on issues of social responsibility and sustainable development, both within organizations themselves and in their relationship with the broader social and natural environment. In this context, the economic, environmental, and social impacts of business operations can no longer be disregarded, regardless of the size, activity, and business model of an organization (Senadheera et al., 2022). In response to growing global demands and international business practices, companies are increasingly taking a proactive approach to managing issues related to environmental protection (*Environment – E*), social responsibility (*Social – S*), and corporate governance (*Governance – G*). The integrated consideration of these aspects of business operations is encompassed by the ESG concept (Passas, 2024), which is increasingly becoming an important framework for assessing the sustainability and responsibility of companies' operations (Li, Wang, Sueyoshi & Wang, 2021).

In recent years, particular attention has been devoted to the environmental component of the ESG concept. This is due to the urgency of the need to reconsider ways of life and draw the world's attention to the accelerated deterioration of the environment and natural resources, as well as to the consequences of such deterioration for economic and social development (Strange & Bayley, 2008). In addition, it is important to note that the environmental component of the ESG concept plays an important role in building trust between shareholders and regulatory authorities, as well as in implementing initiatives aimed at sustainable business practices within companies. This encourages investors to consider sustainable investments, while motivating consumers to choose sustainable products (Mohanty, Priya, Paithe & Mohanty, 2025).

Information on ESG aspects of business operations is presented through ESG reports. Therefore, ESG reporting represents a paradigm shift in corporate reporting, going beyond traditional financial reporting and encompassing a broader range of non-financial information (Aziz & Alshdaifat, 2024; Beck, Dumay & Frost, 2017).

ESG reporting represents an important mechanism for reducing information asymmetry between companies and the market, effectively informing stakeholders, managing risks, reducing costs, improving business performance, strengthening competitive advantage, and enhancing corporate reputation (Jovanović, 2025; Ellili, 2022). Effective and transparent ESG reporting plays a key role in building trust in business operations and, consequently, in companies' resilience to crises and unexpected events, while also representing a fundamental driver of sustainable performance (Jovanović & Jakšić, 2025).

In the preceding period, an increasing number of researchers have examined the relationship between ESG reporting and financial aspects of business operations, with particular attention devoted to examining the relationship between reporting on environmental issues, i.e., environmental performance, and financial performance. The results of previous studies indicate that there is no consensus regarding the nature of this relationship, and the debate concerning their relationship remains relevant today. Empirical analyses of the impact of reporting on the E component on companies' strategies, operations, and financial performance have likewise produced inconclusive and inconsistent results. A review of the existing literature does not provide uniform answers, as, on the one hand, some studies find no relationship (Nor, Bahari, Adnan, Kamal & Ali, 2016; Deswanto & Siregar, 2018; Qiu, Shaukat & Tharyan, 2016; Asuquo, Dada & Onyeogaziri, 2018; Connelly & Limpaphayom, 2004; Malarvizhi & Matta, 2016; Maji & Lohia, 2026), while, on the other hand, some studies indicate the existence of a positive (Karambu & Joseph, 2016; Lu & Taylor, 2018; Haninun, Lindrianasari & Denziana, 2018; Chouaibi, Rossi, Siggia & Chouaibi, 2021; Alfalih, 2023; Chung, Bayne & Birt, 2024; Arshad, Audi & Ali, 2025) or negative (Al-Waeli, Khalid, Ismail & Idan, 2021; Jones, Aruobogha, Mmamefune, Ogheneovo & Eloho, 2024) relationship between reporting on environmental aspects of business operations and companies' financial performance. The results of the studies differ and vary depending on the measures used for reporting, measures of financial performance, measures of environmental performance, sample characteristics, sample size, countries' development status, company size, industry, sector, observation period, and similar factors (Aggarwal, 2013), indicating the need for further and more in-depth analysis of the relationship between the E component of ESG reporting and companies' financial performance.

In addition, an analysis of the existing literature reveals that a large number of existing studies focus on analyzing the aforementioned relationship at the level of individual countries, as well as that ESG is often considered as a unified concept, without separating its components and analyzing the impact of individual ESG components on financial performance. Given that different aspects of companies' operations may have different effects on financial performance, their individual analysis may lead to new and more comprehensive conclusions. Therefore, the main objective of this paper is to determine whether there is a relationship between the E component of ESG reporting, expressed through the environmental component index calculated as the ratio of the total number of disclosed environmental indicators to

the total number of environmental reporting indicators defined in accordance with the GRI Standards, and companies' financial performance, expressed by the return on assets and return on equity.

This research contributes to addressing the identified research gaps. Unlike the predominant focus of previous research on examining the aforementioned relationship at the level of individual countries, this paper focuses on analyzing companies operating in fifteen countries, thereby examining the relationship between the E component of ESG reporting and financial performance in a broader European context. In addition, by focusing exclusively on the environmental component of ESG reporting, this paper complements existing research that considers ESG as a whole and provides additional empirical evidence on its relationship with financial performance. The contribution and novelty of this paper arise from the selected sample. Namely, the research included 100 large companies over a five-year period from 2019 to 2023, out of a total of 227 European companies listed in the publicly available *Carbon Disclosure Project (CDP)* database that were ranked highest according to CDP criteria and received an A score. Descriptive statistics and Spearman's correlation were applied in the empirical part of the research.

The paper is structured into five sections. Following the introduction, the paper presents the theoretical foundations and a review of previous research. The second section presents the research methodology. The third section encompasses the analysis and discussion of the research results, while the final section presents the concluding remarks of the research.

1. Theoretical foundations and review of previous research

1.1. Theoretical foundations

The relationship between ESG reporting and financial performance can be analyzed from several theoretical perspectives, including stakeholder theory, signaling theory, legitimacy theory, and institutional theory.

Stakeholder theory indicates that companies should not focus exclusively on the interests of their shareholders, but also on the interests of other stakeholders who may influence companies or be influenced by them. Accordingly, in addition to owners, i.e., shareholders, companies should also take into account the interests of non-owner stakeholders (Mahmud, 2020). From the perspective of stakeholder theory, the success of companies depends on the successful management of the interests of all stakeholders, mutual trust, and cooperation, without favoring the interests of a particular group (Jovanović, 2026). Based on the foregoing, ESG reporting by companies is considered an issue that is important to a wide range of stakeholders. It encompasses issues that go beyond economic interests, but which may also have an impact on economic aspects of business operations (Lokuwaduge & Heenetigala, 2017; Jenkins, 2004). Definitions of corporate sustainability

explicitly refer to stakeholders (Lokuwaduge & Heenetigala, 2017; Searcy, 2016). Given that sustainability-related decision-making is closely linked to stakeholder needs, the concept of ESG reporting is clearly connected to stakeholder theory, which emphasizes that companies are obliged to meet the needs of stakeholders, both within and outside the company itself (Korankye, Hao, Borah, Odai & Ahakwa, 2024; Lokuwaduge & Heenetigala, 2017; Freeman & McVea, 2005). In fact, stakeholder theory views ESG as a response to the demands of various stakeholders which, when effectively managed, improve company performance through reducing the costs of conflicts, better access to resources, and improved business performance (Salem & Hassan, 2026). In the context of reporting on environmental aspects of business operations, stakeholder theory assumes that information on a company's environmental impact is disclosed due to the information needs of various stakeholders. Particular emphasis should be placed on the numerous stakeholder groups that primarily focus on environmental aspects of business operations. Namely, the long-term survival and success of a company largely depend on the resources provided by external stakeholders; accordingly, stakeholder demands should be taken into account when implementing sustainable and, in particular, environmental activities, as well as when reporting on such activities.

Companies often send signals that reduce information asymmetry between themselves and stakeholders, enabling the transmission of information about their image, intentions, behavior, and performance (López-Santamaría, Amaya, Hinestroza & Cuero, 2021). Regardless of whether a signal is positive or negative, intentional or unintentional, it has the potential to reduce information asymmetry by supplementing the existing body of information available to external stakeholders (Friske, Hoelscher & Nikolov, 2023). In the context of sustainability, company managers (signal senders) use ESG reporting to signal to stakeholders (recipients) their intentions regarding the implementation of sustainable business practices (Friske et al., 2023; Connelly, Ketchen & Slater, 2011). Management uses ESG reports to signal its commitment to environmental protection, social responsibility, and corporate governance, while external stakeholders interpret these signals to assess the performance and long-term prospects of companies (Huang, Hu, Wang & Wang, 2025; Connelly, Certo, Ireland & Reutzel, 2011). Positive signals may contribute to an increase in company value and improved financial performance, while negative signals may adversely affect stakeholders' assessments and attitudes (Kurniawan & Nugroho, 2025). This transparency not only improves long-term corporate performance but also mitigates the risks associated with managers' focus on short-term objectives, which could negatively affect investments in sustainable practices (Deng, Liu, Yan & Ma, 2025).

Legitimacy theory is based on the idea that a company can survive and achieve growth through acceptance by society (Alsayegh, Abdul Rahman & Homayoun, 2020; Guthrie & Parker, 1989). Therefore, legitimacy theory holds that companies should disclose certain information (about involvement in community activities, human resources, physical resources, contribution to environmental protection, etc.)

in order to convince society that the company's activities are acceptable and contribute to the creation of social value. Due to the increasing pressure and attention that stakeholders devote to environmental protection, the disclosure of information on a company's environmental impact and measures to prevent and reduce such impact may provide companies with potential benefits, including increased transparency, employee motivation, access to new sources of financing, as well as improvements in reputation and brand value. In this way, companies may avoid negative market perceptions that may be associated with a reputation for irresponsible treatment of the environment and improve their business performance (Alsayegh et al., 2020; Brammer & Pavelin, 2008). In fact, ESG activities and reporting on environmental protection, social responsibility, and corporate governance contribute to engagement with stakeholders, which may have a positive effect on ESG performance as well as on the company's financial performance (Lee & Raschke, 2023).

Institutional theory is the most commonly used framework in research in the field of ESG reporting because it provides a better understanding of various institutional variables that influence companies' ESG reporting (Jovanović, 2025). Institutional theory suggests that companies are not driven solely by the goal of profit maximization, but also by the influence of various institutions (e.g., institutional investors or governments), reflected in rules, values, beliefs, and assumptions that are taken for granted and accepted as given (Herold, 2018). Institutional pressures and expectations may influence companies to gradually adapt their behavior and business practices to these expectations. Companies may benefit from providing a greater volume of high-quality ESG information, and specifically information on environmental protection, for example through a positive response or receiving recognition from the government and the public, which may have a positive effect on the reputation and "acceptance" of companies, and ultimately on their business performance (Yixi & Sharon, 2023).

The common basis of the aforementioned theoretical perspectives lies in the assumption that ESG reporting can contribute to improving a company's image and reputation, reducing information asymmetry, and improving business performance, thereby creating conditions for long-term and sustainable growth.

1.2. The relationship between the E component of ESG reporting and companies' financial performance

When it comes to the existence of a relationship between the disclosure of information on the environmental aspects of companies' operations and financial performance, the literature points to numerous aspects, therefore, the following section will focus on studies available through Google Scholar, Web of Science, and Scopus that, in the authors' opinion, are relevant considering the subject and objective of the research. Generally speaking, the results of the studies indicate unclear and inconsistent conclusions, as some studies, on the one hand, point to the existence of a positive or negative relationship between reporting on environmental

aspects of business operations and companies' financial performance, while, on the other hand, some studies indicate the absence of such a relationship (Al-Waeli et al., 2021).

One of the more comprehensive studies is the study conducted by Friede, Busch and Bassen (2015). The authors analyzed the results of 2.200 scientific papers dealing with the relationship between sustainability reporting and improvements in financial performance. The results showed that more than 90% of the studies indicated a positive correlation between these two phenomena. Similar findings were reached by Karambu and Joseph (2016), who used a causal research design to determine the cause-and-effect relationship between the E component of ESG reporting and financial performance. The findings revealed a positive relationship between environmental reporting and financial performance. In addition, the study indicated the importance of companies engaging in environmental reporting and managers adopting policies that are aligned with the concept of sustainable development. The objective of the study conducted by Nor et al. (2016) was to determine the existence of a relationship between the E component of ESG reporting and financial performance, using a sample of the 100 largest companies in Malaysia. Content analysis of the annual reports of the observed companies produced mixed results. Namely, the results of this study showed a significant relationship between environmental reporting and profit margin. However, the results obtained did not show a statistically significant relationship between environmental reporting and financial performance, expressed through return on assets and return on equity. The researchers also indicated that, in Malaysia, companies emphasize reporting on their contribution to environmental activities in order to attract investors and meet stakeholder requirements.

Lu and Taylor (2018) indicated that an environmentally responsible company may have a lower value because the immediate benefits of its increased engagement in environmental activities may not become visible immediately, but only five or ten years later. At the same time, an environmentally responsible company may be more inclined to report to the public on its environmental practices in order to demonstrate, through more transparent reporting on its environmental aspects of business operations, that it genuinely cares about society and the community. Examining large U.S. companies, the researchers determined that financial performance and environmental reporting are interrelated, as well as that companies that orient their strategies toward environmental activities generally integrate these activities into all aspects of their operations, which may have long-term effects on their financial success and public perception. Haninun et al. (2018) conducted a study to determine whether environmental performance and the E component of ESG reporting affect companies' financial performance. The results of multiple linear regression showed that environmental performance and environmental reporting, expressed through an environmental reporting index, have a statistically significant positive effect on financial performance, measured by return on assets and return on equity. This was supported by stakeholder theory, legitimacy theory, signaling theory, and political-

economic theory, which indicate that social and environmental reporting in annual reports is a strategic tool for achieving company objectives and a tool that may influence stakeholders' attitudes.

Mixed results regarding the relationship between ESG disclosure, companies' market value, and their performance were also obtained in the study by Giannopoulos, Kihle Fagernes, Elmarzouky and Afzal Hossain (2022), who indicated a positive relationship between ESG scores and company market value, as well as a negative relationship between sustainability reporting and financial profitability, expressed through return on assets (ROA). Similar results were obtained by Al-Waeli et al. (2021), who analyzed a total of 31 studies. The analyses showed mixed results: 19 studies demonstrated a statistically significant positive correlation between environmental reporting and financial performance, while 12 showed a negative correlation in the short term. The researchers indicated that different environmental reporting indicators lead to different effects on different measures of financial performance; nevertheless, most studies indicate that reporting on environmental aspects of business operations improves companies' reputation as well as financial performance. Furthermore, the analysis established that a theoretical framework encompassing legitimacy theory, stakeholder theory, and agency theory provides strong support for ESG reporting. Chouaibi et al. (2021) view the disclosure of information on a company's environmental impact as the latest innovation in the field of corporate reporting. The researchers conducted a study aimed at determining the impact of the E component of ESG reporting on companies' financial performance, while examining the moderating effect of social and ethical practices. The results showed a positive and statistically significant relationship between environmental reporting and financial performance, implying that a high level of disclosure of information on a company's environmental impact contributes to improved financial performance.

Alfalih (2023) conducted an analysis aimed at examining the interaction between the ESG dimensions, i.e., environmental, social, and corporate governance dimensions, and the financial performance of companies included in the S&P 500 index. The results confirmed a relationship between the social and governance dimensions and company performance expressed through return on assets and Tobin's Q. Regarding the environmental dimension, a statistically significant relationship was observed with Tobin's Q. Overall, the results of the study indicated that ESG reporting practices significantly affect companies' financial performance, both directly and indirectly. A similar study was conducted by Chung et al. (2024). The main objective of the study was to determine the impact of disclosure of information on environmental, social, and corporate governance (ESG) issues on companies' financial performance under mandatory reporting conditions in Hong Kong. The researchers established a significant positive relationship between comprehensive ESG reporting and companies' financial performance. Examining the relationship between individual ESG components and companies' financial performance, they established a statistically significant relationship between the E component and

return on assets. In a bibliometric study based on more than 1.000 studies published between 2015 and 2020, Whelan, Atz, Van Holt and Clark (2021) demonstrated the existence of positive, negative, mixed, or no relationships between sustainability reporting and companies' financial success.

Arshad et al. (2025) investigated how the E component of ESG reporting affects the financial performance of companies operating in the energy, mining, manufacturing, and chemical industries. The results of the study showed that a higher level of reporting on the E component is positively and statistically significantly associated with return on assets, return on equity, and Tobin's Q. These results support the stakeholder theory and legitimacy theory perspectives. In addition, the researchers emphasized that proactive and strategic disclosure of information on the E component may contribute to increased profitability, while reactive reporting or reporting primarily motivated by compliance with regulatory requirements may generate additional costs. Maji and Lohia (2026) conducted a study in which they emphasized that ESG reporting contributes to an increase in company value. In addition, the results of their study indicated that, among the three components of ESG reporting, only reporting on corporate governance has a statistically significant effect on company performance. The interaction effect model showed that the relationship between ESG reporting and financial performance is conditioned by the industry in which companies operate, but that this effect is more pronounced in industries that are less sensitive to climate change.

A considerable number of studies have not established a relationship between environmental reporting and financial performance. One of the earliest studies was conducted by Connelly and Limpaphayom (2004). Namely, they conducted a study using a unique dataset from a survey of the Thai Institute of Directors on corporate governance and concluded that there was no statistically significant relationship between environmental reporting and accounting performance. In addition, the researchers suggested that disclosure of good environmental policies does not negatively affect short-term profitability. Similar results were obtained by Malarvizhi and Matta (2016), who examined the relationship between business performance and the level of reporting on the environmental aspect of business operations using a sample of 85 companies from the chemical, energy, and metal sectors. Through content analysis, they found that there was no statistically significant relationship between the level of environmental reporting and companies' business performance. Qiu et al. (2016) likewise did not establish any relationship between the E component of ESG reporting and profitability through their analysis.

Deswanto and Siregar (2018), in their study, suggested that there is no relationship between the E component of ESG reporting and investors' assessment of companies' financial performance, as one type of stakeholder. Asuquo et al. (2018) showed that there is no statistically significant relationship between the E component of ESG reporting and companies' financial performance, expressed through return on assets. Jones et al. (2024) found that there is a negative, but statistically insignificant, relationship between financial performance and the E component of ESG reporting.

The aforementioned research findings indicate that the results of studies are not uniform and that there is no generally accepted view regarding the relationship between reporting on environmental issues and financial performance.

Considering the aforementioned studies that provide contradictory results, the following hypothesis will be tested in this paper:

H1: There is a statistically significant positive relationship between the E component of ESG reporting and companies' financial performance.

The following two sub-hypotheses were developed from the stated hypothesis:

H1a: There is a statistically significant positive relationship between the E component of ESG reporting and return on assets.

H1b: There is a statistically significant positive relationship between the E component of ESG reporting and return on equity.

2. Methodology of the research

In accordance with the defined research objective and research hypotheses, an empirical study was conducted using a quantitative methodology based on statistical data analysis through descriptive statistics and Spearman's correlation coefficient, using the SPSS software package (*Statistical Package for the Social Sciences*). Based on the results obtained, conclusions will be drawn using the methods of analysis, synthesis, and comparison.

The empirical research was conducted on a sample of 100 large European companies included in the publicly available Carbon Disclosure Project (CDP) database, which were ranked highest according to the CDP criteria and received an A score. CDP assesses companies, cities, countries, and regions based on their transparency and the measures they undertake in relation to climate, water, and forests. Although reporting environmental data through CDP is not mandatory, during 2025, approximately 22.100 companies worldwide disclosed data on their environmental impact through CDP. These companies represent more than half of the total global market capitalization and include some of the largest companies in the world, with the most significant environmental impact. At the same time, more than 1.000 cities, countries, and regions, collectively representing more than one billion people, disclosed their data during 2025 (Carbon Disclosure Project, 2025). At the time the data were retrieved for analysis, the CDP database contained a total of 227 European companies that had received an A score. Consequently, the sample comprised 44% of the total number of European companies recorded in the CDP database at the time under observation. The research was conducted by collecting and analyzing secondary data for a five-year period, from 2019 to 2023. The data were collected from available sustainability reports prepared in accordance with the GRI Standards and financial statements (balance sheets and income statements) published on the

websites of the analyzed European companies. The structure of the sample by country is presented in Table 1.

Table 1: Sample structure

Country		Number of companies
	United Kingdom	17
	Spain	15
	France	16
	Norway	4
	Belgium	2
	Switzerland	7
	Turkey	11
	Austria	5
	Sweden	2
	Denmark	1
	Portugal	3
	Germany	9
	Netherlands	3
	Finland	1
	Italy	4
Industry	Manufacturing Industry	51
	Energy Industry	15
	Consulting	7
	Banking and Financial Services	5
	Real Estate	6
	Telecommunications	5
	Retail	3
	Transportation	3
	Marketing	1
	Hospitality	1
	Other	3

Source: Authors

To test the proposed hypothesis, a correlation analysis was conducted. Spearman's correlation coefficient was used to indicate the potential existence of a relationship between the variables relating to the E component of ESG reporting, on the one hand, and financial performance, on the other hand, as well as the strength of the relationship, if one exists. The strength of the correlation was determined according to Cohen, where $r=0.10$ to 0.29 indicates a small correlation, $r=0.30$ to 0.49 indicates a medium correlation, and $r=0.50$ to 1.00 indicates a large correlation (Cohen, 1998).

Financial performance was expressed through profitability in the form of the return on assets (*ROA*) and return on equity (*ROE*). The *ROA* and *ROE* indicators were selected to express financial performance because they have been used as indicators of financial performance in a larger number of empirical studies (Asuquo et al., 2018; Nor et al., 2016; Haninun et al., 2018; Karambu & Joseph, 2016; Hussain, 2015; Alfalih, 2023; Chung et al., 2024; Arshad et al., 2025). The return on assets was measured as the ratio of net income to the average total assets, while the return on equity was measured as the ratio of net income to the average amount of capital employed (Alfalih, 2023; Xu & Wang, 2018; Liu, Zhang, Xu & Wang, 2022). On the other hand, the E component of ESG reporting was expressed through an environmental component index, measured as the ratio of the total number of disclosed environmental indicators to the total number of environmental reporting indicators defined in accordance with the GRI Standards (Hussain, 2015; Hongming et al., 2020; Burhan & Rahmanti, 2012).

3. Analysis and discussion of research results

The proposed hypothesis was tested using Spearman's correlation coefficient.

Table 2 presents the minimum and maximum values, arithmetic means, medians, and standard deviations of the dependent and independent variables.

Table 2: Minimum and maximum values, arithmetic means, medians, and standard deviations of the dependent and independent variables

	Arithmetic mean	Standard deviations	Minimum	Maximum	Median
Environmental component of ESG reporting	0.5880	0.21343	0.24	0.92	0.5800
ROA	0.2310	0.65624	-0.03	5.25	0.1420
ROE	0.3310	1.28310	-0.13	10.13	0.1680

Source: Authors

Based on Table 2, the results of the descriptive statistical analysis for the observed five-year period can be seen. The average value of the environmental component reporting index is 0.59, indicating that the companies under observation, on average, disclose more than half of the indicators for reporting on the environmental aspects of business, compared to the total expected number of indicators, in accordance with the GRI Standards. The obtained average value indicates that the analyzed companies have a relatively high level of disclosure of information on the environmental aspects of business, although there are differences among companies in the extent of disclosed information. The results are expected, considering that the sample consists of companies that voluntarily disclosed data on the impact of their

business activities on the environment through CDP, representing a significant share of the largest European companies, with the most significant impact on the environment. Analyzing the average values of the profitability indicators, positive average values of both indicators can be observed, indicating that the companies included in the sample, on average, generated a positive return on assets and return on equity during the observed period, i.e., that the observed companies in Europe predominantly operated profitably during the period from 2019 to 2023.

Table 3 presents the results of the correlation analysis. Spearman's correlation coefficient was used to analyze the relationship between the variables under consideration.

Table 3: Correlation between variables

Variables		Environmental component of ESG reporting	ROA	ROE
Environmental component of ESG reporting	Correlation coefficient	1	0.104	0.117
ROA	Correlation coefficient	0.104	1	0.983**
ROE	Correlation coefficient	0.117	0.983**	1

**Correlation is significant at the 0.01 level.

Source: Authors

Based on Table 3, it can be concluded that there is a positive and very weak correlation between the E component of ESG reporting and financial performance, as measured by return on assets and return on equity. However, considering that the obtained correlations are not statistically significant, it can be concluded that, in the observed sample, no statistically significant relationship was established between the environmental component of ESG reporting and the financial performance of companies, which is similar to the results obtained by Nor et al. (2016), Qui et al. (2016), Jones et al. (2024), Connelly and Limpaphayom (2004), Malarvizhi and Matta (2016), and Maji and Lohia (2026). Based on the obtained results, it is concluded that the proposed hypothesis and sub-hypotheses cannot be accepted. The analysis also reveals a statistically significant positive and strong correlation between the financial performance indicators, indicating their relationship in the observed sample. The results of the present study differ from the results obtained by Lu and Taylor (2018), Haninun et al. (2018), Al-Waeli et al. (2021), Karambu and Joseph (2016), Chouaibi et al. (2021), Alfalih (2023), Chung et al. (2024), and Arshad et al. (2025), whose findings point to the existence of a relationship between

the environmental component of ESG reporting and the financial performance of companies.

A greater number of arguments may contribute to explaining the obtained results. Some of them generally relate to the relationship between ESG reporting and financial performance, while others can be more directly linked to the relationship between the environmental component of ESG reporting and the financial performance of companies.

One of the causes of the obtained results may be found in the level of environmental costs arising from various environmental initiatives. These costs are most often treated as operating expenses, which may be associated with lower company profitability, particularly in the short term. This aspect is particularly relevant given the specific characteristics of the companies included in the observed sample and the fact that they are required to address environmental protection issues and report on emissions. In this regard, Asuquo et al. (2018) particularly emphasized the significance of environmental costs.

Another reason may be found in the existence of a time lag between the implementation of ESG initiatives, their disclosure, and the realization of potential effects on the financial performance of companies (Lehenchuk, 2023). The benefits that may arise from improving ESG practices, such as enhancing the company's image and reputation, increasing customer loyalty, and improving environmental performance, may not be immediately reflected in financial indicators but may only materialize after a certain period of time. Lehenchuk (2023), analyzing the relationship between ESG reporting and financial performance, specifically points to the importance of observing the effects of ESG activities over the long term, considering that their contribution to company value may not become evident in the short term.

Third, the obtained results may be explained by the heterogeneity of the regulatory environment in the field of ESG reporting in the countries from which the companies included in the sample originate. Namely, during the observed period, there were significant differences in terms of the scope and nature of companies' obligations regarding the disclosure of sustainability information, whereby ESG reporting was not mandatory for certain companies in all countries and in the same manner. Such regulatory heterogeneity may contribute to differences in the scope, quality, and comparability of disclosed ESG information, thereby making it more difficult to determine its relationship with financial performance at the level of the entire sample. This aspect is particularly significant given that the research was conducted on a sample of companies for the period from 2019 to 2023, i.e., before the commencement of the application of the new requirements of Directive (EU) 2022/2464 on corporate sustainability reporting (CSRD). During this period, the regulatory framework for sustainability reporting was not yet harmonized to the extent envisaged by the new European regulatory framework. Therefore, it is possible that differences in regulatory requirements, as well as in the level of

development of ESG reporting practices among countries, represent one of the factors contributing to the statistically insignificant relationship between the environmental component of ESG reporting and the financial performance of companies.

Fourth, the reasons should also be sought in the sectoral characteristics of the sample, given that some sectors are under greater pressure from regulations and legislation to address environmental aspects of business, environmental practices, and reporting. Similar arguments are presented in their study by Maji and Lohia (2026). Therefore, a significant scope of environmental and ESG reporting as a whole does not automatically imply the existence of a relationship with financial results. The analyzed sample is dominated by manufacturing companies, accounting for 66%, from different sectors, which, due to their more direct impact on the environment, are more exposed to stricter regulatory requirements. Thus, the mere presence of environmental practices and reporting on them does not automatically ensure positive financial results.

Fifth, it is not sufficient for companies to have a partial or narrow focus solely on the E component of the ESG concept. Informed and interested stakeholders require a holistic approach to the ESG concept and reporting, as well as the simultaneous implementation of various ESG practices over a longer period. Furthermore, formalized environmental risk management may have financial relevance (Čupić, 2026). Reporting on the company's impacts on the environment should not be viewed as an isolated variable, but rather as part of a broader mechanism through which a company builds trust, reduces uncertainty for investors, and achieves better results.

Sixth, some authors explain the absence of a statistically significant relationship between reporting on the environmental component and the financial performance of companies by inadequate and non-credible reporting (Lehenchuk, 2023). Regulatory requirements and stakeholder pressures may encourage greenwashing as a form of fraud (Vuković, Todorović & Jovanović, 2025). Jovanović (2026) also discusses the significant risk of the impact of greenwashing and stakeholder perceptions in his study. As a result, such reports may be subject to distortions and selective presentation of information, which may not accurately reflect the actual impacts of a company on the environment. Furthermore, the financial relevance of ESG practices depends not only on their existence, but also on adequate communication and the extent to which they provide reliable information on risks and the quality of management (Čupić, 2026). Since the present study did not focus on the quality of reporting, it is not possible to make explicit claims in this regard; however, future research should also include this additional dimension.

Overall, the relationship between environmental reporting and financial performance is neither simple nor one-dimensional. ESG activities and practices may contribute to improving financial performance, but their relationship, as well as their potential causality, depends on the specific ESG practices, the characteristics of the reporting

itself, the time horizon of observation, and the broader context in which the company operates. Finally, the absence of unambiguous conclusions in previous research provides a basis for conducting new studies aimed at gaining a better understanding of the relationship between the environmental component of ESG reporting and the financial performance of companies.

Conclusion

The main objective of the research conducted in this paper was to determine whether there is a relationship between the E component of ESG reporting, expressed through the environmental component index calculated as the ratio of the total number of disclosed environmental indicators to the total number of environmental reporting indicators defined in accordance with the GRI Standards, and the financial performance of companies, expressed through return on assets and return on equity. The results indicated that there is no statistically significant relationship between the E component of ESG reporting, expressed through the environmental component index, and the financial performance of companies, expressed through return on assets and return on equity. The obtained results may be explained by the fact that the relationship between ESG reporting and companies' financial performance generally does not manifest itself immediately, but becomes visible over a longer period of time, through strengthening reputation, customer loyalty, and improving relationships with stakeholders. The quality of ESG reports and sincerity in reporting represent key factors for achieving a genuine relationship, while formal reporting (Lehenchuk, 2023) and the use of greenwashing in reporting are not associated with significant financial results. In the long term, transparent and credible reporting on sustainable practices may be positively related to financial performance, while the short-term absence of a statistically significant relationship is often associated with insufficiently developed strategies, low-quality ESG reports, and companies' insufficient familiarity with sustainable activities (Firmansyah, Umar & Jibril, 2023). Sector-specific characteristics and regulatory pressures on the analyzed sample of companies may also represent causes of the obtained results. Socoliuc et al. (2025) similarly emphasize that the obtained results may be a consequence of differences among the sectors in which companies operate, inconsistent reporting of ESG information by companies, as well as the context in which the research was conducted.

Given the relevance of the topic of ESG reporting and its relationship with companies' financial performance at the global level, this paper contributes to the advancement of the existing literature and the expansion of knowledge in this field. In particular, this is because it includes a relevant sample of companies included in the CDP database, one of the significant international databases on companies' environmental impact. The credibility of the database is reflected in its global coverage and the large number of companies that voluntarily disclose data on their environmental impacts and that encompass a significant share of the world's largest companies, with the most significant impact on the environment. In addition, this

paper also has certain practical implications. The research results may suggest to managers that successful long-term business operations require comprehensive, strategic thinking and action, rather than simply focusing on specific aspects of the ESG concept in the short term. Furthermore, the paper raises the issue of the credibility and quality of ESG reports and draws attention to the fact that the mere existence of a report, i.e., compliance with formal requirements, is not sufficient. Mere compliance with formal requirements and companies' desire to present themselves as better than they actually are create a risk of greenwashing, meaning that genuine financial and other effects are not realized. Partial commitment and public declarations without substantive and comprehensive ESG transformation are not sufficient. From the perspective of the general public and numerous stakeholders, it may, in a certain sense, be fair that sustainability and environmental protection should indeed come at the "expense of capital and returns."

The research conducted in this paper has certain limitations that may provide a useful basis and direction for future research in this field. The limitations identified during the analysis and processing of data include the fact that the analysis covered only the first 100 European companies on the Carbon Disclosure Project list (out of a total of 227 European companies on the aforementioned list), that only a five-year period was analyzed, that there may be a risk regarding the comparability of the reports under observation, and that financial performance was expressed only through the ROA and ROE indicators. Future research should include a larger sample, the use of multiple measures of financial and environmental performance through several different indicators, different categories of environmental costs, different aspects of report quality, and a longer period of analysis.

References

- Aggarwal, P. (2013). Sustainability reporting and its impact on corporate financial performance: A literature review. *Indian Journal of Commerce and Management Studies*, 4(3), 51-59.
- Alfalih, A. A. (2023). ESG disclosure practices and financial performance: a general and sector analysis of SP-500 non-financial companies and the moderating effect of economic conditions. *Journal of Sustainable Finance & Investment*, 13(4), 1506-1533. <https://doi.org/10.1080/20430795.2022.2150511>
- Alsayegh, M. F., Abdul Rahman, R., & Homayoun, S. (2020). Corporate economic, environmental, and social sustainability performance transformation through ESG disclosure. *Sustainability*, 12(9), 3910. <https://doi.org/10.3390/su12093910>
- Al-Waeli, A. J., Khalid, A. A., Ismail, Z., & Idan, H. Z. (2021). The Relationship between environmental disclosure and financial performance of industrial companies with using a new theory: Literature Review. *Journal of Contemporary Issues in Business and Government*, 27(2), 3846-3868. 10.47750/cibg.2021.27.02.393

Arshad, R., Audi, M., & Ali, A. (2025). Environmental disclosure and financial performance: Evidence from environmentally sensitive sectors across global markets. *Policy Journal of Social Science Review*, 3(8), 383-399. <https://www.policyjssr.com/index.php/PJSSR/article/view/438>

Asuquo, A. I., Dada, E. T., Onyeogaziri, U. R. (2018). The effect of sustainability reporting on corporate performance of selected quoted brewery firms in Nigeria. *International Journal of Business & Law Research*, 6(3), 1-10.

Aziz, N. H. A., & Alshdaifat, S. (2024). ESG Reporting: Impacts, Benefits and Challenges. In: Alshurafat, H., Hamdan, A., Sands, J. (eds.). *Sustainable Horizons for Business, Education, and Technology. Contributions to Environmental Sciences & Innovative Business Technology* (69-76). Singapore: Springer. https://doi.org/10.1007/978-981-97-2981-4_5.

Beck, C., Dumay, J., & Frost, G. (2017). In pursuit of a ‘single source of truth’: from threatened legitimacy to integrated reporting. *Journal of Business Ethics*, 141(1), 191-205. <https://doi.org/10.1007/s10551-014-2423-1>

Brammer, S., & Pavelin, S. (2008). Factors influencing the quality of corporate environmental disclosure. *Business Strategy and the Environment*, 17(2), 120-136. <https://doi.org/10.1002/bse.506>

Burhan, A. H. N., Rahmanti, W. (2012). The impact of sustainability reporting on company performance. *Journal of Economics, Business, and Accountancy Ventura*, 15(2), 257-272.

Carbon Disclosure Project. (2025). *CDP Scores and A Lists*, retrieved on September 5, 2026, from: <https://www.cdp.net/en/data/scores>

Chouaibi, S., Rossi, M., Siggia, D., & Chouaibi, J. (2021). Exploring the moderating role of social and ethical practices in the relationship between environmental disclosure and financial performance: Evidence from ESG companies. *Sustainability*, 14(1), 209. <https://doi.org/10.3390/su14010209>

Chung, R., Bayne, L., & Birt, J. (2024). The impact of environmental, social and governance (ESG) disclosure on firm financial performance: evidence from Hong Kong. *Asian Review of Accounting*, 32(1), 136-165. <https://doi.org/10.1108/ARA-07-2022-0165>

Cohen, J. W. (1998). *Statistical power analysis for the behavioral sciences* (2nd ed.). NJ: Lawrence Erlbaum Associates.

Connelly, B. L., Certo, S. T., Ireland, R. D., & Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, 37(1), 39-67. <https://doi.org/10.1177/0149206310388419>

Connelly, B. L., Ketchen, D. J., & Slater, S. F. (2011). Toward a “theoretical toolbox” for sustainability research in marketing. *Journal of the Academy of Marketing Science*, 39(1), 86–100. <https://doi.org/10.1007/s11747-010-0199-0>

Connelly, J. T., Limpaphayom, P. (2004). Environmental reporting and firm performance: evidence from Thailand. *Journal of Corporate Citizenship*, 13, 137-149.

Čupić, M. (2026). ESG prakse kao determinanta dostupnosti izvora finansiranja za kompanije u Srbiji, U: Mikerević, D. (ur.). *Evolucija računovodstvene i revizorske profesije* (375-402). Banja Luka: SRRRS.

Deng, G., Liu, H., Yan, J., & Ma, S. (2025). Managing for the future: managerial short-termism impact on corporate ESG performance in China. *The European Journal of Finance*, 31(2), 147-173. <https://doi.org/10.1080/1351847X.2024.2387622>

Deswanto, R. B., & Siregar, S. V. (2018). The associations between environmental disclosures with financial performance, environmental performance, and firm value. *Social Responsibility Journal*, 14(1), 180-193. <https://doi.org/10.1108/SRJ-01-2017-0005>

Ellili, N. O. D. (2022). Impact of ESG disclosure and financial reporting quality on investment efficiency. *Corporate Governance: The International Journal of Business in Society*, 22(5), 1094-1111. <https://doi.org/10.1108/CG-06-2021-0209>

Firmansyah, E. A., Umar, U. H., & Jibril, R. S. (2023). Investigating the effect of ESG disclosure on firm performance: The case of Saudi Arabian listed firms. *Cogent Economics & Finance*, 11(2), 2287923. <https://doi.org/10.1080/23322039.2023.2287923>

Freeman, R. E., & McVea, J. (2005). A Stakeholder Approach to Strategic Management. *The Blackwell handbook of strategic management*, 183-201. <https://doi.org/10.1111/b.9780631218616.2006.00007.x>

Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210-233.

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>

Giannopoulos, G., Kihle Fagernes, R. V., Elmarzouky, M., & Afzal Hossain, K. A. B. M. (2022). The ESG disclosure and the financial performance of Norwegian listed firms. *Journal of Risk and Financial Management*, 15(6), 237. <https://doi.org/10.3390/jrfm15060237>

Guthrie, J., & Parker, L. D. (1989). Corporate social reporting: a rebuttal of legitimacy theory. *Accounting and Business Research*, 19(76), 343-352. <https://doi.org/10.1080/00014788.1989.9728863>

- Haninun, H., Lindrianasari, L., & Denziana, A. (2018). The effect of environmental performance and disclosure on financial performance. *International Journal of Trade and Global Markets*, 11(1-2), 138-148. <https://doi.org/10.1504/IJTGM.2018.092471>
- Herold, M. H. (2018). Demystifying the link between institutional theory and stakeholder theory in sustainability reporting. *Economics, Management and Sustainability*, 3(2), 6-19. <https://doi.org/10.14254/jems.2018.3-2.1>
- Hongming, X., Ahmed, B., Hussain, A., Rehman, A., Ullah, I., & Khan, F. U. (2020). Sustainability reporting and firm performance: The demonstration of Pakistani firms. *Sage Open*, 10(3), 1-12. <https://doi.org/10.1177/2158244020953180>
- Huang, J., Hu, P., Wang, D. D., & Wang, Y. (2025). The double signal of ESG reports: Readability, growth, and institutional influence on firm value. *Sustainability*, 17(6), 2514. <https://doi.org/10.3390/su17062514>
- Hussain, N. (2015). *Impact of sustainability performance on financial performance: an empirical study of global fortune (N100) firms*, retrieved on September 5, 2026, from: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2616835
- Jenkins, H. (2004). Corporate social responsibility and the mining industry: conflicts and constructs. *Corporate Social Responsibility and Environmental Management*, 11(1), 23–34 <https://doi.org/10.1002/csr.50>
- Jones, A. S., Aruobogha, O. E., Mmamefune, O. L., Ogheneovo, O. C., & Eloho, O. O. (2024). Corporate sustainability reporting and financial performance of likelihood distress companies in Nigeria. *International Journal of Applied Economics, Finance and Accounting*, 20(1), 78-88. <https://doi.org/10.33094/ijaefa.v20i1.1926>
- Jovanović, D. (2025). Izazovi i prepreke ESG izveštavanja u Republici Srbiji. 19. međunarodni simpozijum o korporativnom upravljanju: održivi razvoj i porezi - šansa ili barijera razvoja malih ekonomija (97-113). Banja Luka: Finrar d.o.o.
- Jovanović, D. (2026). *Sistemi upravljanja zaštitom životne sredine – od održivog razvoja do izveštavanja o održivosti*. Kragujevac: Ekonomski fakultet Univerziteta u Kragujevcu.
- Jovanović, D., Jakšić, M. (2025). Ekološka dimenzija izveštavanja o održivosti – normativni okvir i analiza prakse u srpskim kompanijama. U: Mikerević, D. (ur.). 29. Međunarodni kongres računovodstvene i revizorske profesije Republike Srpske – Savremeni računovodstveni izazovi: multidimenzionalna priroda složenih informacionih potreba (201-222). Banja Vrućica.
- Karambu, K. G., & Joseph, M. W. (2016). Effect of Corporate Environmental Disclosure on Financial Performance of Firms Listed at Nairobi Securities Exchange, Kenya. *International Journal of Sustainability Management and Information Technologies*, 2(1), 1-6. 10.11648/j.ijsm.20160201.11

- Korankye, B., Hao, Y., Borah, P. S., Odai, L. A., & Ahakwa, I. (2024). Transformational leadership, ESG performance, corporate reputation and competitive advantage: a serial mediation model. *Business Process Management Journal*, 31(6), 2182–2203. <https://doi.org/10.1108/BPMJ-08-2024-0692>
- Kurniawan, F., & Nugroho, A. H. L. (2025). ESG performance during financial performance and reporting quality shortfalls: Proving signaling theory in Indonesia. *The Indonesian Accounting Review*, 15(1), 99-110. 10.14414/tiar.v15i1.5246
- Lee, M. T., & Raschke, R. L. (2023). Stakeholder legitimacy in firm greening and financial performance: What about greenwashing temptations?. *Journal of Business Research*, 155, 113393. <https://doi.org/10.1016/j.jbusres.2022.113393>
- Lehenchuk, S., Zhyhlei, I., Ivashko, O., & Gliszczynski, G. (2023). The impact of sustainability reporting on financial performance: Evidence from Turkish FBT and TCL Sectors. *Sustainability*, 15(20), 14707. <https://doi.org/10.3390/su152014707>
- Li, T. T., Wang, K., Sueyoshi, T., & Wang, D. D. (2021). ESG: Research progress and future prospects. *Sustainability*, 13(21), 11663. <https://doi.org/10.3390/su132111663>
- Liu, L., Zhang, J., Xu, J., & Wang, Y. (2022). Intellectual capital and financial performance of Chinese manufacturing SMEs: an analysis from the perspective of different industry types. *Sustainability*, 14(17), 10657. <https://doi.org/10.3390/su141710657>
- Lokuwaduge, C. S. D. S., & Heenetigala, K. (2017). Integrating environmental, social and governance (ESG) disclosure for a sustainable development: An Australian study. *Business Strategy and the Environment*, 26(4), 438-450. <https://doi.org/10.1002/bse.1927>
- López-Santamaría, M., Amaya, N., Hinestroza, M. P. G., & Cuero, Y. A. (2021). Sustainability disclosure practices as seen through the lens of the signaling theory: A study of companies listed on the Colombian Stock Exchange. *Journal of cleaner production*, 317, 128416. <https://doi.org/10.1016/j.jclepro.2021.128416>
- Lu, L. W., & Taylor, M. E. (2018). A study of the relationships among environmental performance, environmental disclosure, and financial performance. *Asian Review of Accounting*, 26(1), 107-130. <https://doi.org/10.1108/ARA-01-2016-0010>
- Mahmud, M. T. (2020). Quest for a single theory to explain managerial motivations for sustainability disclosures: Legitimacy theory, stakeholder theory or institutional Theory. *Bulletin of Japanese Association for International Accounting Studies*, 7, 135-159.
- Maji, S. G., & Lohia, P. (2026). Unveiling the financial effect of ESG disclosure on financial performance in India: climate-sensitive corporates' perspective. *International Journal of Ethics and Systems*, 42(1), 109-131. <https://doi.org/10.1108/IJOES-02-2024-0039>

- Malarvizhi, P., Matta, R. (2016). Link between corporate environmental disclosure and firm performance—perception or reality?. *Review of Integrative Business & Economics Research*, 5(3), 1-34.
- Mohanty, A., Priya, V. K., Paithe, U. N., & Mohanty, A. (2025). 'E' of Environmental, Social, and Governance (ESG) in India: reporting, review and future prospects. *Discover Global Society*, 3(1), 136. <https://doi.org/10.1007/s44282-025-00288-0>
- Nor, N. M., Bahari, N. A. S., Adnan, N. A., Kamal, S. M. Q. A. S., & Ali, I. M. (2016). The effects of environmental disclosure on financial performance in Malaysia. *Procedia Economics and Finance*, 35, 117-126. [https://doi.org/10.1016/S2212-5671\(16\)00016-2](https://doi.org/10.1016/S2212-5671(16)00016-2)
- Passas, I. (2024). The evolution of ESG: From CSR to ESG 2.0. *Encyclopedia*, 4(4), 1711-1720. <https://doi.org/10.3390/encyclopedia4040112>
- Qiu, Y., Shaukat, A., & Tharyan, R. (2016). Environmental and social disclosures: Link with corporate financial performance. *The British Accounting Review*, 48(1), 102-116. <https://doi.org/10.1016/j.bar.2014.10.007>
- Salem, M. R., & Hassan, R. (2026). Do sustainability theories hold empirically? Evidence from a theory-driven systematic review of ESG and sustainability finance. *Environment Systems and Decisions*, 46(4), 82. <https://doi.org/10.1007/s10669-026-10144-4>
- Searcy, C. (2016). Measuring enterprise sustainability. *Business Strategy and the Environment*, 25(2), 120–133. <https://doi.org/10.1002/bse.1861>
- Senadheera, S. S., Gregory, R., Rinklebe, J., Farrukh, M., Rhee, J. H., & Ok, Y. S. (2022). The development of research on environmental, social, and governance (ESG): A bibliometric analysis. *Sustainable Environment*, 8(1), 2125869. <https://doi.org/10.1080/27658511.2022.2125869>
- Socoliuc, M. I., Ionescu, B. Ș., Guinea, F. A., Socoliuc, M., Grosu, V., & Hlaciuc, E. (2025). The evaluation of the relationship between the companies' sustainability reporting and their financial performance. *Amfiteatru Economic*, 27(70), 884–902. <https://doi.org/10.24818/EA/2022/5946>
- Strange, T., & Bayley, A. (2008). *Sustainable Development: Linking Economy, Society, Environment*. Paris: OECD Publishing.
- Vuković, I., Todorović, M., Jovanović, D. (2025). Greenwashing as a form of fraud in sustainability reports. In: Malinić, D., Vučković Milutinović, S. (edi.). *Second Conference on Forensic Accounting: Contemporary forensic accounting trends in combating financial fraud: global and national perspectives* (195-209). Belgrade: Faculty of Economics and Business University of Belgrade.
- Whelan, T., Atz, U., Van Holt, T., Clark, C. (2021). ESG and financial performance. *Uncovering the relationship by aggregating evidence from*, 1(1), 2015-2020.

Xu, J., & Wang, B. (2018). Intellectual capital, financial performance and companies' sustainable growth: Evidence from the Korean manufacturing industry. *Sustainability*, 10(12), 4651. <https://doi.org/10.3390/su10124651>

Yixi, Z., & Sharon, C. P. Y. (2023). The influence of firm size and institutional environment on ESG disclosure-evidence from listed companies in China. *International Journal of Academic Research in Economics and Management and Sciences*, 12(2), 109-132. 10.6007/IJAREMS/v12-i2/16284

Amira Pobrić
University of East Sarajevo –
Faculty of Economics Brčko
ORCID: 0000-0001-8902-7223

IDENTIFYING COMPANIES WITH AN INCREASED RISK OF FRAUDULENT FINANCIAL REPORTING USING ACCOUNTING VARIABLES

UDK 657.6:343.53(497.6)

DOI <https://doi.org/10.54318/czf.2026.ap.017>

Abstract

Existing models for detecting fraudulent financial reporting have largely been developed and validated on samples of companies from developed economies. Consequently, their applicability in developing countries may be limited by differences in institutional and economic environments. The limited availability of data on confirmed fraudulent financial reporting cases constrains the adaptation of existing models and the development of new models tailored to developing countries. In light of these limitations, the aim of this study is to identify companies with increased risk of fraudulent financial reporting using selected accounting variables derived from the published financial statements of large non-financial companies in Bosnia and Herzegovina.

The study was conducted on a sample of 382 companies using six accounting variables that prior research has identified as being significantly associated with confirmed fraudulent financial reporting and as proxies for financial pressure, opportunities for fraud, or the potential exploitation of such opportunities. High-risk values for each accounting variable were determined using quartile and decile thresholds. Companies with an increased risk of fraudulent financial reporting were identified based on the number, persistence, and type of accounting variables with high-risk values.

The results show that, using quartile thresholds, 73 companies with high-risk values observed for four or more accounting variables were identified as companies with an increased risk of fraudulent financial reporting. Within this group, narrower subsets of companies with particularly pronounced accounting risk signals were distinguished based on the persistence of high-risk values and the application of stricter decile thresholds. Analysis of the types of accounting variables for which high-risk values occurred simultaneously, together with their persistence, enabled the identification of an additional 26 companies with an increased risk among those for which high-risk values were observed for two or three accounting variables. The findings indicate that jointly considering the number, persistence, and type of accounting variables with high-risk values enables the identification of companies that may require more detailed forensic and audit examination.

Key words: fraudulent financial reporting; accounting variables; fraud risk assessment; large non-financial companies; Bosnia and Herzegovina

JEL Classification: M41, K13

Introduction

Financial statements are the primary source of information for investors, creditors, regulators, and other stakeholders in making business decisions. Their usefulness depends on their ability to faithfully represent a company's financial position, financial performance, and cash flows. However, in practice, financial statements do not always reflect the economic reality of a company. Experiences from numerous financial scandals have shown that management may engage in fraudulent financial reporting in order to present a more favorable picture of a company's financial position and business performance, most often under pressure to achieve planned results, obtain external sources of financing, or meet the expectations of investors and creditors (Kanapickienė & Grundienė, 2015; Malinić & Ribić, 2025; Narsa, Afifa & Wardhaningrum, 2023).

Fraud is defined as any intentional act involving deception, concealment, or abuse of trust to obtain an unjust or unlawful benefit (The Institute of Internal Auditors, 2024, p. 11). In the context of financial reporting, it most commonly takes the form of the intentional misstatement or omission of material information intended to mislead users of financial statements (Association of Certified Fraud Examiners, 2026). Unlike earnings management, which may occur within the boundaries of accounting standards, fraudulent financial reporting involves their deliberate violation (Perols & Lougee, 2011). The consequences of such behaviour extend beyond individual companies, undermining confidence in financial reporting and capital markets and leading to substantial investor losses, regulatory interventions, and the inefficient allocation of economic resources (Beneish, Lee & Nichols, 2013; Malinić & Ribić, 2025).

The timely detection of fraudulent financial reporting is an important element in protecting users of financial statements. Since fraud generally cannot be established solely on the basis of publicly available information, fraudulent financial reporting risk assessment is an important step in identifying companies that require more detailed forensic, audit, or regulatory investigation. In this context, financial statements may have considerable diagnostic value because they constitute a standardized and widely available source of information from which unusual financial reporting patterns can be identified (Milićević & Janjić, 2025).

No single accounting variable can, in itself, constitute evidence of fraudulent financial reporting. However, certain patterns in financial statements may indicate an increased likelihood of opportunistic accounting behaviour (Beneish, 1999; Dechow, Ge, Larson & Sloan, 2011). Accordingly, numerous models and approaches have been developed to detect manipulation and material misstatements in financial statement, among which the Beneish M-score (Beneish, 1999) and the Dechow F-score (Dechow et al., 2011) are the most widely recognised, which combine multiple accounting variables into a single measure of the risk. These models rely on predetermined coefficients of accounting variables estimated using samples of companies from developed economies, which raises the question of their

direct applicability in different institutional and economic environments. Differences in regulatory frameworks, the quality of institutional mechanisms, and financial reporting practices may affect their predictive ability and indicate the need to verify their applicability and possible adaptation in specific institutional conditions (Marais, Vermaak & Shewell, 2023). However, the adaptation of existing models is often limited by the lack of publicly available data on confirmed cases of fraudulent financial reporting in developing countries, which directs researchers to explore alternative approaches for identifying companies with an increased risk of fraudulent financial reporting.

Given the limited availability of data on confirmed cases of fraudulent financial reporting in Bosnia and Herzegovina and the limited applicability of existing models developed for different institutional settings, the aim of this study is to identify companies with an increased risk of fraudulent financial reporting using accounting variables derived from the published financial statements of large non-financial companies in Bosnia and Herzegovina. The analysis is based on accounting variables that prior empirical research has identified as being significantly associated with confirmed cases of fraudulent financial reporting and material misstatements in financial statements (Beneish, 1999; Dechow et al., 2011). The study contributes to the development of approaches for identifying companies with an increased risk of fraudulent financial reporting and provides a practical basis for identifying companies that require more detailed forensic investigations.

The remainder of the paper is organised as follows. Section 2 reviews the theoretical foundations of fraudulent financial reporting and systemizes empirical evidence on the relationship between accounting variables and confirmed cases of fraudulent financial reporting. Section 3 describes the research methodology. Section 4 presents the empirical findings, while Section 5 concludes the paper.

1. Literature review

Theoretical explanations of fraudulent financial reporting are largely based on agency theory and fraud theories, which seek to explain the motives, circumstances, and individual characteristics that lead to fraudulent behaviour. Agency theory (Jensen & Meckling, 1976) is based on the assumption that the separation of ownership and control creates a conflict of interest between shareholders and management. The separation of ownership and control also creates information asymmetry, allowing managers to take actions that are not necessarily aligned with the interests of shareholders, but are directed towards the realization of personal interests. When corporate governance and monitoring mechanisms are ineffective, managers can exploit information asymmetry for opportunistic behavior, including fraudulent financial reporting.

The most influential theoretical framework for explaining the causes of fraud is the fraud triangle theory (Cressey, 1953), which posits that fraud occurs when pressure, opportunity, and rationalisation are simultaneously present. Pressure refers to

financial or other circumstances that motivate individuals to engage in fraudulent behaviour; opportunity refers to the ability to commit and conceal fraud arising from weaknesses in internal control and monitoring mechanisms; and rationalisation allows perpetrators to justify their fraudulent actions as acceptable (Dorminey, Fleming, Kranacher & Riley, 2012). Later research has expanded this theoretical framework by incorporating additional factors. Wolfe and Hermanson (2004) developed the diamond theory of fraud by adding capability as a fourth element that enables committing and concealing fraud. The fraud pentagon theory further incorporates arrogance (Marks, 2014), whereas the fraud hexagon theory introduces collusion as an additional element associated with more complex forms of fraudulent behaviour (Vousinas, 2019). Although these models differ in the number of elements they include, they all seek to explain the circumstances under which individuals and organisations engage in fraudulent behaviour.

Because rationalisation and perpetrators' individual characteristics cannot be operationalised using information contained in financial statements, research based on accounting information has primarily focused on indicators of financial pressure and accounting items that, owing to the significant degree of subjective assessments involved, may provide opportunities for fraudulent financial reporting (Beneish, 1999).

Numerous empirical studies confirm that information contained in financial statements can assist in identifying companies with an increased risk of fraudulent financial reporting. Summers and Sweeney (1998) demonstrated that certain accounting variables, established on the basis of financial statements, enable distinguishing companies involved in fraudulent financial reporting from other companies. Beneish (1999) developed a model based on accounting variables that distinguishes earnings manipulators from non-manipulators, while Dechow et al. (2011) showed that combining accounting variables and market variables significantly improves the identification of material misstatements in financial statements. These findings indicate that information contained in financial statements does not, in itself, constitute evidence of fraud, but may contain signals that enable timely fraud risk assessment and the identification of companies with an increased likelihood of fraudulent financial reporting.

Within the fraud triangle theory, financial pressure is considered one of the primary drivers of fraudulent behaviour. Management may be motivated to conceal poor operating performance, meet investors' and creditors' expectations, or avoid the consequences of failing to comply with contractual obligations. Empirical evidence shows that financial pressure may be reflected in various financial characteristics of a company, such as low profitability, high growth rate, increased need for external sources of financing, and high level of financial indebtedness (Dechow et al., 2011; Lou & Wang, 2009; Skousen, Smith & Wright, 2009).

With regard to profitability, the literature generally assumes that weak financial performance increases the likelihood that management will resort to more aggressive

accounting practices in order to present more favorable financial performance. This assumption is confirmed by Spathis (2002), who has found that companies involved in fraudulent financial reporting are characterized by lower profitability, while Beneish (1999) found that deterioration in gross margins was positively and statistically significantly associated with the likelihood of earnings manipulation. In contrast, Dalnial, Kamaluddin, Sanusi & Khairuddin (2014) and Narsa et al. (2023) did not find a statistically significant association between profitability and fraudulent financial reporting.

Companies that achieve high growth rates are often under pressure to maintain the expected business dynamics. Beneish (1999) found that rapid sales growth is associated with financial statement manipulation, whereas Skousen et al. (2009) found that higher asset growth is associated with a greater likelihood of fraudulent financial reporting. By contrast, Narsa et al. (2023) found a negative association between asset growth and fraudulent financial reporting, suggesting that companies with lower asset growth face greater pressure to manipulate financial results in order to present a more stable financial position.

Companies with insufficient internally generated cash flows are often forced to obtain additional sources of financing, which may create an incentive to report more favorable financial results. Consistent with this argument, Skousen et al. (2009) found that increased cash and external financing needs are significantly associated with fraudulent financial reporting. In contrast, Dechow et al. (2011) did not find a statistically significant association between the need for external financing and material accounting misstatements.

Empirical evidence generally confirms that higher financial leverage is associated with a greater likelihood of fraudulent financial reporting (Dalnial et al., 2014; Lou & Wang, 2009; Narsa et al., 2023; Spathis, 2002). This relationship is explained by the fact that highly leveraged companies face a greater risk of breaching debt covenants and failing to meet obligations to creditors, thereby increasing management's incentive to satisfy the expectations of creditors and other external stakeholders (Lou & Wang, 2009). However, Beneish (1999) found that leverage did not significantly contribute to distinguishing earnings manipulators from non-manipulators, stating that incentives to comply with debt covenant restrictions may not be strong enough to induce earnings manipulation, because the costs of violating them may be relatively low.

In addition to financial pressure, the fraud triangle theory identifies opportunity as one of the key conditions for committing fraud. In the context of financial reporting, opportunities for manipulation are particularly pronounced in accounting areas that require a significant degree of professional judgment or involve subjective accounting estimates. For example, measuring accounts receivable requires estimates of collectability, while inventory valuation often involves estimates of net realizable value, obsolescence, and cost allocation. The limited ability to independently verify reported values in these areas increases the scope for financial

statement manipulation. Empirical research shows that fictitious, unrealized, or prematurely recognized revenues, overstated inventories, and improper capitalization of costs are among the most common forms of material misstatement in financial statements (Beneish, 1999; Dechow et al., 2011; Loebbecke, Eining & Willingham, 1989). Based on these findings, numerous empirical studies have identified accounting variables that can serve either as proxies for opportunities to engage in fraudulent financial reporting or as signals that such opportunities have been exploited.

Accounting variables that reflect changes in the proportion of assets whose measurement involves substantial subjective management judgement may serve as proxies for opportunities to engage in fraudulent financial reporting. The underlying premise is that an increase in the proportion of such assets expands the scope for the exercise of accounting discretion, thereby increasing management's ability to influence the recognition and measurement of individual financial statement items and, consequently, the reported financial position and financial performance of the company. Empirical confirmation of this assumption is provided by Beneish (1999) and Dechow et al. (2011). Beneish (1999) showed that such indicators help distinguish companies that manipulate earnings from those that do not, whereas Dechow et al. (2011) found that they are significant predictors of material misstatements in financial statements.

Accounts receivable represent one of the accounting areas in which management exercises substantial accounting discretion, which may be exploited to manipulate financial reporting. An unusual increase in accounts receivable may result from legitimate business reasons, such as changes in credit policy aimed at stimulating sales or changes in the composition of the customer base. However, an unusual increase in accounts receivable may indicate premature revenue recognition, the recording of fictitious sales, or insufficient recognition of impairment losses. Consequently, variables based on accounts receivable are often used as signals of the potential exploitation of opportunities for fraudulent financial reporting. Most empirical studies have found that accounts receivable-based variables play a significant role in identifying irregularities in financial reporting (Beneish, 1999; Dalnial et al., 2014; Dechow et al., 2011), although some studies have not found them to be statistically significant (Skousen et al., 2009).

Inventories are among the most significant asset categories in many companies and have a direct effect on both financial position and financial performance. Consequently, even relatively small misstatements in inventory valuation may materially affect reported earnings and assets. At the same time, inventory measurement requires considerable management judgement in estimating net realisable value, inventory obsolescence and the allocation of production costs, creating opportunities for the manipulation of reported financial results. Accordingly, unusually high or otherwise abnormal values of inventory-related variables may indicate the potential exploitation of such opportunities, particularly through inventory overstatement resulting from improper cost capitalisation or other

accounting practices that delay expense recognition. Most empirical research confirms the usefulness of inventory-based variables in identifying companies involved in fraudulent financial reporting (Dalnial et al., 2014; Spathis, 2002; Summers & Sweeney, 1998), although Skousen et al. (2009) did not find this relationship to be statistically significant.

Unlike accounting variables related to individual accounting areas, accruals reflect the cumulative effects of management's accounting estimates and discretionary decisions. Since accruals arise from the application of the accrual basis of accounting and incorporate numerous estimates and assumptions regarding future economic benefits and obligations, they may provide a useful signal that management's accounting discretion has been exploited in financial reporting. Empirical evidence generally confirms that total accruals are associated with an increased risk of fraudulent financial reporting and material misstatements (Beneish, 1999; Dechow et al., 2011). However, more sophisticated measures of accruals, based on the assessment of their unexpected or discretionary part, do not necessarily provide better results in identifying fraudulent financial reporting than simpler measures of total accruals (Bayley & Taylor, 2007; Jones, Krishnan & Melendrez, 2008).

A review of the existing literature suggests that fraudulent financial reporting is a complex phenomenon arising from the interaction of various forms of financial pressure and opportunities for manipulation. Empirical evidence indicates that some accounting variables may serve as proxies for financial pressure or opportunities for manipulation, whereas others may signal the potential exploitation of such opportunities. Although many accounting variables have been found to be associated with fraudulent financial reporting, the reported findings are not entirely consistent. Since fraudulent financial reporting is influenced by multiple risk factors, prior research generally combines accounting variables reflecting different dimensions of fraud risk rather than relying on individual variables.

2. Research methodology

The empirical analysis is based on the calculation and analysis of selected accounting variables for companies in Bosnia and Herzegovina. The sample consists of large non-financial companies in Bosnia and Herzegovina for which all data required to calculate the selected accounting variables were available. Large companies were selected because they are characterised by more complex corporate governance structures and because financial reporting in such companies is particularly demanding in terms of both scale and sophistication, making them an appropriate setting for investigating the risk of fraudulent financial reporting (Malinić & Ribić, 2025). Companies operating in the financial sector were excluded because of the specific nature of their business activities, regulatory framework, and financial reporting practices.

The sample design was determined by the fact that different legal criteria for classifying companies by size were applied in the Federation of Bosnia and

Herzegovina and the Republic of Srpska during the observed period (2023–2025). Consequently, the sample was constructed using the company size classification criteria applicable in the Federation of Bosnia and Herzegovina (Zakon o računovodstvu i reviziji u Federaciji Bosne i Hercegovine, 2021), which have also been applied in the Republic of Srpska since 2026 (Zakon o računovodstvu i reviziji Republike Srpske, 2025). Because the total number of companies in the Republic of Srpska meeting these criteria could not be determined, the corresponding population size could not be established. The final sample comprises 382 large companies, including 288 from the Federation of Bosnia and Herzegovina and 94 from the Republic of Srpska.

Data for the analysis were obtained from the annual financial statements of the sampled companies for the period 2023–2025. The use of three consecutive years enabled the calculation of accounting variables based on year-to-year changes and the examination of whether high-risk values occurred in only one or both of the observed annual comparisons.

The data were obtained from the Financial Information Agency of the Federation of Bosnia and Herzegovina and the Agency for Intermediary, IT and Financial Services of the Republic of Srpska. Although different financial statement formats are used in the two entities, resulting in differences in the classification and coding of certain financial statement items, all accounting variables were constructed using equivalent accounting items, thus ensuring data comparability.

Based on the literature review, six accounting variables identified in prior research as being associated with fraudulent financial reporting were selected. The research variables are presented in Table 1. The operationalisation of the research variables was adapted to the objective of the study and the research design.

The total accruals variable (ACC) was constructed using the same approach as the total accruals measure in the Beneish model, differing only in the denominator used for scaling. Whereas Beneish (1999) scales total accruals by total assets at the end of the period, this study scales them by average total assets, which reduces potential distortions that can arise due to significant changes in the size of the company over the accounting period.

The variables $\Delta(\text{AR}/\text{SALES})$ and $\Delta(\text{INV}/\text{SALES})$ were operationalised as the absolute change in the accounts receivable-to-sales ratio and the inventory-to-sales ratio, respectively, between two consecutive years. Unlike the corresponding variable in the Beneish model, which is based on the relative change in the accounts receivable-to-sales ratio, the variable in this study is based on the absolute change in that ratio. This approach avoids overstating changes in companies with low initial values of the ratio and is therefore more appropriate for identifying high-risk values of accounting variables. The same approach was applied to the operationalisation of $\Delta(\text{INV}/\text{SALES})$, ensuring methodological consistency in the operationalisation of both variables.

Table 1: Selected accounting variables

Variable	Variable name	Formula
ACC	Total accruals	$\Delta \text{Current Assets} - \Delta \text{Cash} - (\Delta \text{Current Liabilities} - \Delta \text{Short-Term Liabilities})$
$\Delta(\text{AR}/\text{SALES})$	Change in the accounts receivable-to-sales ratio	$\left(\frac{\text{Accounts Receivable}_t}{\text{Sales}_t} \right) - \left(\frac{\text{Accounts Receivable}_{t-1}}{\text{Sales}_{t-1}} \right)$
$\Delta(\text{INV}/\text{SALES})$	Change in the inventory-to-sales ratio	$\left(\frac{\text{Inventory}_t}{\text{Sales}_t} \right) - \left(\frac{\text{Inventory}_{t-1}}{\text{Sales}_{t-1}} \right)$
LEV	Financial leverage	$\frac{\text{Short-term Financial Liabilities} + \text{Long-term Financial Liabilities}}{\text{Total Assets}}$
SOFT	Soft assets ratio	$\frac{\text{Total Assets} - \text{Cash} - \text{Property, Plant and Equipment}}{\text{Total Assets}}$
ΔROA	Change in return on assets	$\left(\frac{\text{Net Income}_t}{\text{Total Assets}_t} \right) - \left(\frac{\text{Net Income}_{t-1}}{\text{Total Assets}_{t-1}} \right)$

Source: Authors' own compilation.

A small number of extreme values were identified for the $\Delta(\text{AR}/\text{SALES})$ and $\Delta(\text{INV}/\text{SALES})$ variables during the analysis. Upon detailed examination, it was determined that some of the extreme values were the result of extremely low sales revenues as the denominator in the formula for calculating the variables, which is why they primarily reflected a very low level of business activity, rather than an increased risk of financial statement manipulation. Such observations were excluded from the analysis, while all others were retained.

The LEV variable was operationalised as the ratio of total financial liabilities to total assets. Unlike the indicator of change in leverage used in some models for detecting financial statement manipulation (Beneish, 1999), this study uses the level of financial leverage, based on the assumption that financial pressure arises primarily from a company's degree of indebtedness rather than from its increase compared to the previous period. The ΔROA variable was defined as the change in return on assets between two consecutive years and was used as an indicator of changes in companies' financial performance.

The SOFT variable was operationalised in accordance with the approach proposed by Dechow et al. (2011) and was defined as the proportion of assets whose valuation depends to a greater extent on management's accounting estimates and assumptions. Although the measurement of property, plant and equipment involves a certain degree of accounting estimation, these assets were excluded from the calculation of the SOFT variable because they provide considerably less scope for discretionary accounting estimates aimed at managing short-term financial performance than other asset categories.

Since there are no universally accepted threshold values for the analysed accounting variables that could reliably indicate an increased risk of fraudulent financial reporting across different institutional environments, this study adopts a relative approach based on the distribution of the values of the accounting variables in the observed sample. The use of distribution-based measures (quartiles, quintiles, and deciles) to define threshold values represents a common methodological approach in empirical accounting and finance research (e.g., Beneish et al., 2013; Bissessur & Veenman, 2016; Frankel & Lee, 1998). In this study, quartiles and deciles were used to define the threshold values. Quartiles were used to identify a broader group of potentially high-risk companies, whereas deciles were used to identify companies with the most extreme values of the accounting variables.

This study is based on the assumption that high-risk values of individual accounting variables do not, by themselves, provide a sufficient basis for identifying companies with an increased risk of fraudulent financial reporting. Therefore, an approach based on three complementary characteristics of the identified accounting risk signals was applied: the number of accounting variables with high-risk values, the persistence of these values over time, and the type of accounting variables for which high-risk values were identified. The number of variables with high-risk values was used to capture the concentration of accounting risk signals, their persistence to determine

whether the same signals recurred in both annual comparisons, while the analysis of variable types enabled consideration of the nature of the identified signals. The application of stricter decile thresholds provided additional information on the extremity of the identified high-risk values.

Based on the abovementioned approach, no single risk score was calculated, nor was a model developed to estimate the probability of fraudulent financial reporting for individual companies. Instead, the abovementioned information was used to identify groups of companies with an increased risk of fraudulent financial reporting that may require more detailed forensic and audit examination.

3. Results and discussion

This section presents the empirical findings obtained by applying the methodology described in the previous section. Table 2 presents the descriptive statistics of the analysed accounting variables in the observed sample. Based on the distribution of their values, quartile and decile thresholds were determined to identify high-risk values. For the variables ACC, $\Delta(\text{AR}/\text{SALES})$, $\Delta(\text{INV}/\text{SALES})$, SOFT, and LEV, higher values indicate an increased risk of fraudulent financial reporting, which is why values above the third quartile and the ninth decile were classified as high-risk values. In contrast, for variable ΔROA , lower values indicate greater financial pressure, so values below the first quartile and the first decile were classified as high-risk values. The application of these thresholds makes it possible to identify companies with the most extreme values of the analysed accounting variables relative to the sample.

Descriptive statistics indicate that the analysed accounting variables differ in terms of their value ranges and the extent of extreme values. For the variables ACC, $\Delta(\text{AR}/\text{SALES})$, $\Delta(\text{INV}/\text{SALES})$, ΔROA , and LEV, the maximum values (or, in the case of ΔROA , the minimum value) are substantially higher than the corresponding quartile and decile thresholds, indicating the presence of highly extreme values. In contrast, the distribution of the SOFT variable is considerably more compact. The difference between the third quartile, the ninth decile, and the maximum value is substantially smaller, which is expected given that the values of this variable are limited by the interval from 0 to 1. In the case of the ΔROA variable, the negative value of the first quartile indicates that companies experiencing a decline in profitability relative to the previous year were classified as high-risk. In the case of the LEV variable, the maximum value exceeding 2 indicates the presence of companies with an exceptionally high degree of financial leverage.

By applying quartile and decile thresholds, 25% and 10% of companies, respectively, were identified for which the high-risk values of individual accounting variables were identified in each annual comparison. To assess the concentration of the identified accounting risk signals at the individual company level, the number of accounting variables for which high-risk values occurred simultaneously was determined for each company in the sample. The results are shown in Table 3.

Table 2: Descriptive statistics and classification thresholds for the analysed accounting variables

Variable	Min	D1	Q1	Median	Q3	D9	Max
ACC	-0.95106	-0.14095	-0.07452	-0.01767	0.03546	0.10648	0.43034
Δ(AR/SALES)	-2.25768	-0.04520	-0.01203	0.00043	0.01864	0.06989	0.65363
Δ(INV/SALES)	-3.82947	-0.03745	-0.00894	0.00048	0.01705	0.04486	1.84080
ΔROA	-4.03344	-0.06018	-0.02239	-0.00166	0.01697	0.05031	0.35413
SOFT	0.00984	0.21512	0.39279	0.57758	0.76237	0.87466	0.99963
LEV	0	0	0.02603	0.16920	0.35497	0.51193	2.12715

Note: D1 and D9 denote the first and ninth deciles of the variable distributions, whereas Q1 and Q3 denote the first and third quartiles.

Source: Author’s research

Table 3: Distribution of companies by the number of simultaneously identified high-risk values of accounting variables

Number of simultaneously identified high-risk values of accounting variables	Quartile thresholds		Decile thresholds	
	Number of companies	Percentage (%)	Number of companies	Percentage (%)
0	30	7.85	148	38.74
1	90	23.56	141	36.91
2	106	27.75	57	14.92
3	83	21.73	24	6.28
4	51	13.35	9	2.36
5	17	4.45	3	0.79
6	5	1.31	0	0.00

Source: Author’s research

The results show substantial differences among companies in the number of accounting variables for which high-risk values are simultaneously present. Using quartile thresholds, no accounting variable with a high-risk value was identified for 30 companies, while for the remaining companies the number ranged from one to six. High-risk values were identified for three accounting variables in 83 companies and for four or more variables in 73 companies, including 5 companies for which high-risk values were identified for all six variables. A greater number of accounting variables with simultaneously present high-risk values indicates a higher concentration of accounting signals associated with fraudulent financial reporting. For companies with high-risk values simultaneously identified for four or more variables, at least one variable must indicate potential accounting consequences of fraudulent financial reporting, which is not necessarily the case for companies with high-risk values simultaneously identified for fewer than four variables. Therefore, the latter group of 73 companies can be identified as a group with an increased risk of fraudulent financial reporting.

Using the decile thresholds, the number of accounting variables with simultaneously present high-risk values was, as expected, lower. High-risk values were identified for three accounting variables in 24 companies and for four or more variables in 12 companies. These 12 companies belong to the previously identified group of 73 companies with an increased risk of fraudulent financial reporting. However, the simultaneous presence of high-risk values for at least four accounting variables identified using the stricter decile thresholds further distinguishes these companies within that group, indicating a particularly pronounced risk of fraudulent financial reporting.

The number of accounting variables with high-risk values and the application of differently strict thresholds enable companies to be differentiated according to the concentration and severity of the identified accounting risk signals, but do not capture their persistence over time. The presence of a high-risk value for the same accounting variable in both annual comparisons indicates the persistence of the corresponding risk signal. Therefore, the persistence of high-risk values was further analysed to distinguish, within the group of 73 companies with an increased risk, those with a particularly pronounced risk of fraudulent financial reporting. The results are presented in Table 4.

The results presented in Table 4 show that, for a large number of the companies analysed, no high-risk value persisted in both annual comparisons for any of the accounting variables, while for a substantial number of companies, a high-risk value persisted for only one variable. On the other hand, the persistence of high-risk values for two or more accounting variables was observed among companies with different total numbers of variables with high-risk values. Of particular relevance are companies in which a larger number of accounting variables with high-risk values is accompanied by the persistence of a larger number of these values.

Table 4: Distribution of companies by number of accounting variables with risk values and their persistence

Panel A: Quartile thresholds							
Number of simultaneously identified high-risk values of accounting variables	Number of variables for which high-risk values retained in both annual comparisons						
	0	1	2	3	4	5	6
1	69	21	-	-	-	-	-
2	50	49	7	-	-	-	-
3	29	30	20	4	-	-	-
4	9	21	16	4	1	-	-
5	5	4	7	1	0	0	
6	0	0	2	1	1	1	0
Panel B: Decile thresholds							
Number of simultaneously identified high-risk values of accounting variables	Number of variables for which high-risk values retained in both annual comparisons						
	0	1	2	3	4	5	6
1	107	34	-	-	-	-	-
2	32	23	2	-	-	-	-
3	16	4	4	0	-	-	-
4	6	2	1	0	0	-	-
5	0	2	0	1	0	0	-
6	0	0	0	0	0	0	0

Source: Author's research

Using quartile thresholds, 9 companies had high-risk values for four or more accounting variables, of which at least three persisted in both annual comparisons. One company was particularly notable, even when the decile thresholds were applied, high-risk values were identified for five accounting variables, three of which persisted in both annual comparisons. Such a combination of the concentration and persistence of accounting risk signals indicates a particularly pronounced risk of fraudulent financial reporting. Three of these nine companies had already been identified as having a particularly pronounced risk because high-risk values for four or more accounting variables were also identified using the decile thresholds. The analysis of persistence therefore enabled the identification of 6 additional companies with a particularly pronounced risk that had not been identified using this stricter criterion for high-risk values.

It should be noted that a larger number of accounting variables with high-risk values is not always accompanied by a larger number of persistent high-risk values. For some companies with four, five, or six variables with high-risk values, persistence was observed for only a smaller number of variables, while for some companies no persistent high-risk values were identified. These results show that the number and persistence of high-risk values provide different but complementary information.

However, the persistence of high-risk values for only a smaller number of variables does not diminish the significance of the simultaneous presence of high-risk values for a larger number of accounting variables.

Companies with the same number of accounting variables with high-risk values may also differ in terms of the types of accounting risk signals that occur simultaneously. This distinction may be particularly important for companies for which high-risk values are identified for two or three accounting variables, as the number of such variables alone provides limited information about the nature of the identified risk. Therefore, the structure of accounting variables for which high-risk values were simultaneously identified was further analysed. The variables were classified according to whether they may indicate financial pressure (Δ ROA and LEV), an opportunity to commit fraud (SOFT), or potential accounting consequences of fraudulent financial reporting (ACC, Δ (AR/SALES), and Δ (INV/SALES)). The results are presented in Table 5.

Table 5: Distribution of companies by number and structure of accounting variables with risk values

Panel A: Quartile thresholds						
Number of simultaneously identified high-risk values of accounting variables	The structure of accounting variables with simultaneously present high-risk values					
	FP	FP+O	AC	O+AC	FP+AC	FP+O+AC
2	7	6	30	12	51	-
3	-	0	8	13	42	20
4	-	-	-	3	19	29
5	-	-	-	-	4	13
6	-	-	-	-	-	5
Panel B: Decile thresholds						
Number of simultaneously identified high-risk values of accounting variables	The structure of accounting variables with simultaneously present high-risk values					
	FP	FP+O	AC	O+AC	FP+AC	FP+O+AC
2	5	8	18	11	15	-
3	-	0	6	0	14	4
4	-	-	-	2	2	5
5	-	-	-	-	0	3
6	-	-	-	-	-	0

Note: FP – variables indicating financial pressure; O – variable indicating an opportunity to commit fraud; AC – variables that may indicate potential accounting consequences of fraudulent financial reporting.

Source: Author's research

Using quartile thresholds, among companies for which high-risk values were identified for two or three accounting variables, the most common combinations comprise financial pressure variables and variables that may indicate potential accounting consequences of fraudulent financial reporting. These are followed by combinations consisting exclusively of variables that may indicate potential accounting consequences of fraudulent financial reporting, as well as combinations of these variables with the variable indicating an opportunity to commit fraud. Similar findings were obtained using the decile thresholds, with expected differences in the absolute frequency of individual combinations of variables. These results show that, for most of these companies, the accounting variables with high-risk values include at least one variable that indicates potential accounting consequences of fraudulent financial reporting. The presence of such variables, either alone or in combination with financial pressure and opportunity variables, provides additional information for identifying companies with an increased risk of fraudulent financial reporting.

A more detailed analysis of this group of companies showed that, among companies for which high-risk values were identified for three accounting variables, 23 companies had high-risk values for two or all three variables that persisted in both annual comparisons. For 11 of these 23 companies, high-risk values for two or all three variables were also identified using the decile thresholds. In addition, among companies for which high-risk values were identified for two accounting variables, 3 companies had high-risk values for both variables that persisted in both annual comparisons, while for 1 of these 3 companies, high-risk values for both variables were also identified using the decile thresholds. The simultaneous presence of variables that may indicate potential accounting consequences of fraudulent financial reporting and the persistence of at least two identified high-risk values provide an additional basis for identifying these companies as having an increased risk of fraudulent financial reporting. Moreover, the identification of high-risk values for at least two accounting variables using the stricter decile thresholds for some of these companies provides further evidence of the pronounced nature of the identified accounting risk signals.

Although the results presented in Table 5 provide an additional basis for identifying companies that could be subject to more detailed forensic examination, the structure of the identified accounting risk signals should be interpreted with caution. The presence of financial pressure variables alone, as well as their combination with the opportunity variable, indicates circumstances that may be conducive to fraudulent financial reporting, but without signals from the set of variables used that may indicate potential accounting consequences of fraudulent financial reporting. The absence of such signals does not exclude the possibility of fraudulent financial reporting, since the set of variables used does not capture all possible forms of manipulation or all of their accounting consequences. Similarly, the absence of identified financial pressure signals does not exclude its existence, as it may be associated with factors not captured by the variables used. Furthermore, fraudulent

financial reporting may affect the values of certain variables used to measure financial pressure, which may result in the absence of high-risk values for these variables. Finally, high-risk values of variables that may indicate potential accounting consequences of fraudulent financial reporting do not constitute evidence that fraud has occurred. Such values may also arise from legitimate business activities and operating conditions and should therefore be regarded as accounting risk signals that require further investigation rather than as confirmation of fraudulent financial reporting.

Conclusion

The aim of this study was to identify companies with an increased risk of fraudulent financial reporting using selected accounting variables. These companies were identified based on the number, persistence, and type of accounting variables with high-risk values. The application of the proposed approach enabled the identification of several groups of companies with an increased risk of fraudulent financial reporting. Using quartile thresholds, 73 companies with high-risk values observed for four or more accounting variables were identified. The simultaneous presence of high-risk values for at least four of the six analysed accounting variables indicates a high concentration of accounting risk signals and provides a basis for identifying these companies as having an increased risk of fraudulent financial reporting. Within this group, 12 companies were further distinguished because high-risk values for at least four variables were also identified using the stricter decile thresholds, as well as 6 companies for which at least three high-risk values persisted in both annual comparisons. These two subgroups may be considered to have a particularly pronounced risk of fraudulent financial reporting.

The analysis of the structure of accounting risk signals provided an additional basis for identifying companies at risk, particularly among companies for which high-risk values were identified for two or three accounting variables. For most of these companies, the identified signals included at least one variable that may indicate potential accounting consequence of fraudulent financial reporting, either alone or in combination with variables indicating financial pressure or an opportunity to commit fraud. Among these companies, 23 companies with three accounting variables with high-risk values were identified, for which at least two high-risk values persisted in both annual comparisons, as well as 3 companies with two accounting variables with high-risk values, for which both high-risk values persisted. For some of these companies, the high-risk values of the same variables were also identified using the stricter decile thresholds, further indicating an increased risk of fraudulent financial reporting.

The main limitation of this study stems from the lack of data on confirmed cases of fraudulent financial reporting, which makes it impossible to verify whether the companies identified as those with increased risk had actually engaged in fraudulent financial reporting. Once data on confirmed cases of fraudulent financial reporting

become available, it will be possible to empirically validate the effectiveness of the proposed approach in identifying companies with an increased risk of fraudulent financial reporting and, based on the findings, further refine the approach.

Furthermore, the study is based on a selected set of accounting variables that have been identified in previous empirical research as significant indicators of fraud risk. The selected variables do not capture all possible indicators of financial pressure or all potential accounting consequences of financial statement manipulation. Accordingly, future research could expand the set of variables used by incorporating additional accounting variables, as well as non-accounting variables related to corporate governance characteristics, external audit characteristics, and similar factors.

References

- Association of Certified Fraud Examiners. (2026). *Occupational fraud 2026: A report to the nations*. retrieved on July 17, 2026, from: <https://www.acfe.com/-/media/files/acfe/pdfs/rtn/2026/2026-report-to-the-nations.pdf>
- Bayley, L., Taylor, S. L. (2007). Identifying earnings overstatements: A practical test. ABN Amro Sydney and University of New South Wales Working paper. retrieved on July 9, 2026, from: https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID995957_code1819.pdf?abstractid=995957&mirid=1
- Beneish, M. D. (1999). The detection of earnings manipulation. *Financial analysts journal*, 55(5), 24-36.
- Beneish, M. D., Lee, C. M., & Nichols, D. C. (2013). Earnings manipulation and expected returns. *Financial Analysts Journal*, 69(2), 57-82.
- Bissessur, S. W., Veenman, D. (2016). Analyst information precision and small earnings surprises. *Review of Accounting Studies*, 21(4), 1327-1360. <https://doi.org/10.1007/s11142-016-9370-2>
- Cressey, D. 1953. *Other People's Money; a Study in the Social Psychology of Embezzlement*. Free Press.
- Dalnial, H., Kamaluddin, A., Sanusi, Z. M., Khairuddin, K. S. (2014). Detecting fraudulent financial reporting through financial statement analysis. *Journal of Advanced Management Science*, 2(1), 17-22. doi: 10.12720/joams.2.1.17-22
- Dechow, P. M., Ge, W., Larson, C. R., Sloan, R. G. (2011). Predicting material accounting misstatements. *Contemporary accounting research*, 28(1), 17-82. <https://doi.org/10.1111/j.1911-3846.2010.01041.x>
- Dorminey, J., Fleming, A. S., Kranacher, M. J., Riley Jr, R. A. (2012). The evolution of fraud theory. *Issues in accounting education*, 27(2), 555-579. <https://doi.org/10.2308/iace-50131>

- Frankel, R., Lee, C. M. (1998). Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and economics*, 25(3), 283-319. [https://doi.org/10.1016/S0165-4101\(98\)00026-3](https://doi.org/10.1016/S0165-4101(98)00026-3)
- Jensen, M., Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs, and ownership structure. *Journal of Financial Economics*, 3(4), 305-360.
- Jones, K., Krishnan, G. V., Melendrez, K. (2007). Do models of discretionary accruals detect actual cases of fraudulent and restated earnings? An empirical evaluation. *Contemporary Accounting Research*, 25(2), 499-531. <https://doi.org/10.57912/23890887>
- Kanapickienė, R., Grundienė, Ž. (2015). The model of fraud detection in financial statements by means of financial ratios. *Procedia-Social and Behavioral Sciences*, 213, 321-327. <https://doi.org/10.1016/j.sbspro.2015.11.545>
- Loebbecke, J. K., Eining, M. M., Willingham, J. J. (1989). Auditors' Experience with Material Irregularities: Frequency, Nature, and Detectability. *Auditing: A journal of practice & theory*, 9(1), 1-28. <https://doi.org/10.2308/AJPT-4682433>
- Lou, Y. I., Wang, M. L. (2009). Fraud risk factor of the fraud triangle assessing the likelihood of fraudulent financial reporting. *Journal of Business & Economics Research*, 7(2), 61-78.
- Narsa, N. P. D. R. H., Afifa, L. M. E., Wardhaningrum, O. A. (2023). Fraud triangle and earnings management based on the modified M-score: A study on manufacturing company in Indonesia. *Heliyon*, 9(2), 36-49. <https://doi.org/10.1016/j.heliyon.2023.e13649>
- Malinić, D., Ribić, M., (2025). Identifying indicators of fraudulent financial reporting under conditions of institutional opacity: The relevance of the Beneish model. In: Malinić, D., Vučković Milutinović, S. (Eds.). *Contemporary forensic accounting trends in combating financial fraud: Global and National Perspectives*. (pp. 1-22). University of Belgrade – Faculty of economics and business.
- Marais, A., Vermaak, C., Shewell, P. (2023). Predicting financial statement manipulation in South Africa: A comparison of the Beneish and Dechow models. *Cogent Economics & Finance*, 11(1), 1-13. <https://doi.org/10.1080/23322039.2023.2190215>
- Marks, J. T. (2014). Playing offense in a high-risk environment: A sophisticated approach to fighting fraud. Crowe Horwath LLP.
- Milićević, V., Janjić, J., (2025). The role of financial statement analysis in detecting corporate fraud. In: Malinić, D., Vučković Milutinović, S. (Eds.). *Contemporary forensic accounting trends in combating financial fraud: Global and National Perspectives*. (pp. 71-87). University of Belgrade – Faculty of economics and business.

Perols, J. L., Lougee, B. A. (2011). The relation between earnings management and financial statement fraud. *Advances in Accounting*, 27(1), 39-53. doi:10.1016/j.adiac.2010.10.004

Spathis, C. T. (2002). Detecting false financial statements using published data: some evidence from Greece. *Managerial Auditing Journal*, 17(4), 179-191. <https://doi.org/10.1108/02686900210424321>

Skousen, C. J., Smith, K. R., Wright, C. J. (2009). Detecting and predicting financial statement fraud: The effectiveness of the fraud triangle and SAS No. 99. Advances in Financial Economics. In: Hirschey, M., John, K., Makhija, A. K. (Eds). *Advances in Financial Economics: Corporate Governance and Firm Performance*. (pp. 53-81). Emerald Group Publishing Limited.

Summers, S. L., Sweeney, J. T. (1998). Fraudulently misstated financial statements and insider trading: An empirical analysis. *Accounting Review*, 73(1), 131-146.

The Institute of Internal Auditors. (2024). *Global internal audit standards*. retrieved on July 16, 2026, from: https://www.theiia.org/globalassets/site/standards/globalinternalauditstandards_2024january9_printable.pdf

Vousinas, G. L. (2019). Advancing theory of fraud: the SCORE model. *Journal of Financial Crime*, 26(1), 372-381. <https://doi.org/10.1108/JFC-12-2017-0128>

Wolfe, D. T., Hermanson, D. R. (2004). The fraud diamond: Considering the four elements of fraud. *The CPA Journal*, 74(12), 38-42.

Zakon o računovodstvu i reviziji Republike Srpske. (2025). Službeni glasnik Republike Srpske br. 115/25.

Zakon o računovodstvu i reviziji u Federaciji Bosne i Hercegovine. (2021). Službene novine FBiH br. 15/2021.

Marina Vasić
University of Belgrade – Faculty of
Agriculture
ORCID: 0000-0002-7926-0053

FORENSIC ACCOUNTING AND THE PROTECTION OF MINORITY SHAREHOLDERS IN EMERGING MARKETS

UDK 657.6:347.72.031

DOI <https://doi.org/10.54318/czf.2026.mv.018>

Abstract

This paper examines the potential role and scope of forensic accounting expertise in protecting minority shareholders in Serbia's underdeveloped capital market. The analysis is grounded in the agency conflict between controlling and minority shareholders, which may intensify in takeover procedures, squeeze-outs, the exercise of dissenting shareholders' rights and disputes over fair compensation.

Although financial reporting regulation, auditing and valuation procedures are intended to ensure the credibility of the information used in such corporate actions, they may not fully eliminate the risks arising from conflicts of interest, information asymmetry and potential financial statement manipulation. Valuations typically rely on company-generated financial data and assumptions that may be influenced by the controlling shareholder, particularly where earnings management, creative accounting or related-party transactions may have affected a company's reported performance and value before the relevant corporate action.

Against this background, the paper explores the potential and limitations of forensic accounting expertise as a mechanism for protecting minority shareholder rights and supporting the fair resolution of valuation and pricing disputes. The analysis also draws on empirical evidence on post-takeover outcomes and selected Serbian disputes. Particular attention is devoted to judicial and regulatory proceedings involving contested valuations and pricing outcomes. By combining valuation expertise with an examination of financial reporting quality, transaction structures, related-party relationships and indicators of possible manipulation, forensic accountants may provide a more independent and comprehensive basis for judicial and regulatory decision-making. The paper argues that forensic accounting experts may contribute to a more reliable determination of fair compensation and an impartial assessment of competing valuation and pricing claims. In doing so, they may also strengthen investor confidence and support the further development of the Serbian capital market.

Key words: forensic accounting; minority shareholder protection; fair compensation; share valuation; emerging capital markets

JEL Classification: M49, G32, K22

Introduction

Emerging capital markets present a specific environment in which limited liquidity, concentrated ownership and information asymmetry may significantly affect the position of outside investors. These conditions may weaken the market mechanisms that ordinarily discipline corporate behavior and protect their interests. Under such circumstances, formal legal rights alone may not be sufficient to ensure effective investor protection. Corporate transparency, the quality and availability of financial information, regulatory enforcement and the practical ability of investors to exercise their rights may be equally important. These concerns become particularly relevant where trading is limited, raising questions about whether market prices provide a reliable indication of underlying economic value.

Serbia provides a relevant setting in this respect. Its stock market was shaped largely by privatization-related listings rather than conventional initial public offerings, with low liquidity and persistent delisting remaining prominent features (Denčić-Mihajlov, 2009). The domestic capital market is characterized by very low investor activity and has limited capacity to fulfil its financing function (Vlada Republike Srbije, 2021). The resulting underdeveloped and illiquid market, with highly concentrated ownership, creates an environment in which conflicts between controlling and minority shareholders may become particularly important. This second agency problem is especially relevant in takeover procedures, subsequent squeeze-outs, the exercise of dissenting shareholders' rights and disputes over fair compensation, where the terms on which minority shareholders may or must exit the company may be difficult to assess or effectively challenge.

Although financial reporting regulation, auditing and valuation procedures are intended to provide credible information and support fair outcomes in such transactions, they may not fully eliminate the risks arising from conflicts of interest and information asymmetry. Valuations typically depend on company-generated financial data, projections and assumptions, while reported performance may itself have been affected by accounting choices and estimates, earnings management, related-party transactions or other decisions preceding the relevant corporate action. In illiquid markets, an additional difficulty arises where market prices may not provide a sufficiently reliable benchmark for fair compensation. Consequently, formal compliance with the applicable price-determination rules, including statutory liquidity criteria, does not necessarily ensure that the resulting price provides a fair and economically reliable basis for minority shareholder exit.

These circumstances create a potential role for forensic accounting expertise that extends beyond the preparation of another valuation. The relevant question is not only whether minority shareholder compensation has been properly determined under the applicable legal and professional standards, but also whether the inputs, underlying financial information and surrounding transactions provide a reliable and economically neutral basis for that determination. This may require examination of financial reporting quality, related-party relationships, ownership and control

structures, pre-transaction trading patterns, valuation assumptions and indicators of possible manipulation. The issue becomes particularly important in judicial proceedings involving conflicting expert valuations and opinions, where understanding why competing assessments differ may be as important as comparing their results.

Against this background, this paper examines the potential and limitations of forensic accounting expertise in protecting minority shareholders in Serbia's underdeveloped capital market. It combines a conceptual examination of situations in which forensic accounting may contribute to minority shareholder protection with empirical evidence on post-takeover outcomes and illustrative evidence from selected Serbian shareholder disputes. Particular attention is devoted to takeovers, squeeze-outs and contested valuation and pricing outcomes, with the aim of assessing how forensic accounting may provide a more independent basis for evaluating the information and transactions underlying minority shareholder compensation and supporting judicial and regulatory decision-making.

1. Minority Shareholder Vulnerability in Key Corporate Events

Minority shareholders are inherently vulnerable because their economic exposure to corporate decisions is not matched by a corresponding ability to influence those decisions or access the information on which they are based. Their protection therefore depends not only on formal legal rights, but also on transparency of control, reliable information, fair compensation and effective remedies where their interests are adversely affected.

1.1. Dissenting Shareholders and Fundamental Corporate Changes

Fundamental corporate changes may expose dissenting shareholders to decisions adopted by the majority that they cannot prevent. Appraisal rights mitigate this conflict by allowing them to exit at fair compensation (Vasiljević & Radović, 2017). Serbian company law grants such rights in relation to certain fundamental decisions, including status changes, major asset transactions and delisting, and allows judicial protection where the relevant value is disputed. The basis for determining compensation depends on whether the shares meet the statutory liquidity criteria: for liquid shares, market value is based on the six-month volume-weighted average price; otherwise, the shareholder is entitled to the higher of book value and appraised value.

1.2. Change of Control through Takeover Bids

A takeover may alter the balance of control and intensify the conflict between controlling and minority shareholders, while the mandatory bid provides remaining shareholders with an exit opportunity at a regulated price and a mechanism for sharing the control premium (Mukwiri, 2013; Mülberr & Sajnovits, 2025). Serbian takeover law generally requires a takeover bid once more than 25% of the voting

shares is acquired, with statutory pricing rules intended to protect remaining shareholders. For liquid shares, the minimum price is the higher of the six-month volume-weighted average market price and the highest price paid by the bidder or persons acting in concert during the preceding 12 months. For non-liquid shares, it is the highest of the preceding acquisition price, book value and appraised fair value. Unlike the EU Takeover Bids Directive, Serbian takeover law extends these pricing constraints to voluntary bids.

Shareholders who consider an approved takeover price inadequate have no separate judicial procedure for determining fair compensation. The Securities Commission verifies compliance with the statutory pricing rules, while judicial review before the Administrative Court concerns legality rather than an independent determination of value. A shareholder may reject the bid but cannot seek an independent judicial determination of a higher price. This is particularly important where the takeover is shortly followed by a squeeze-out on the same pricing terms.

1.3.Squeeze-Out, Sell-Out and Minority Shareholder Exit

Under Serbian company law, squeeze-out and sell-out rights generally arise when a shareholder holds at least 90% of both share capital and the voting rights attached to ordinary shares. Under the ordinary regime, compensation follows the valuation rules applicable to dissenting shareholders and may be judicially contested. Where the 90% threshold deemed to have been reached through a takeover bid, however, squeeze-out and sell-out may be carried out on the terms of that bid for three months after it expires. A further vulnerability arises from the imprecise statutory wording, which leaves room for an expansive interpretation that the special regime can apply even where the bidder acquired only part of its 90% holding through the takeover bid, crossing the threshold by combining a pre-existing stake with shares acquired through the offer. Taken to its extreme, this interpretation would allow the special regime to apply where the bidder crossed the 90% threshold by acquiring only a single share through the takeover bid. The takeover-bid price would then serve as the compensation in the subsequent squeeze-out, even though the remaining minority shareholders had not accepted that price.

Together, these features heighten the vulnerability of minority shareholders, as fair compensation becomes the central protective mechanism in a compulsory exit (Mukwiri, 2013). A shareholder who does not accept the bid may subsequently be compelled to exit at the same price, although its economic adequacy could not be directly challenged during the takeover and the ordinary squeeze-out valuation procedure is unavailable¹.

¹ However, the Court of Justice of the European Union (CJEU), in *Svema Trade* (Case C-567/24, judgment of 27 November 2025, ECLI:EU:C:2025:920), held that the takeover-bid price is only presumed to be fair for a subsequent squeeze-out. That presumption is rebuttable where circumstances call the fairness of the price into question, so a shareholder later compelled to exit at the same price must be able to challenge its fairness under the Directive. This judgement is relevant in light of Serbia's alignment with EU takeover law.

2. Serbia's Stock Market: Post-Takeover Outcomes and Minority Shareholder Exit

To examine takeover characteristics and subsequent outcomes on the Belgrade Stock Exchange (BSE), we conducted an empirical analysis based on data from the Central Securities Depository and Clearing House of the Republic of Serbia (CRHOV), BSE and the Serbian Business Registers Agency (SBRA). As of July 18, 2026, the CRHOV database contained 924 successfully completed takeover bids, from which 280 bids were randomly selected using Excel-generated random numbers; the selected bids related to 280 distinct target companies. BSE and SBRA data were used to identify subsequent delisting, recent trading activity and registration status. To examine recent takeover and pricing patterns, a supplementary analysis covered the period 2021–2026, selected by working backwards approximately five years from the latest available 2026 data. It included 62 takeover bids and 16 amendments published by the BSE to July 26, 2026, relating to 54 target companies.

We found that withdrawal from the stock market was the dominant post-takeover outcome. Only 16 of the 280 target companies (5.7%) remained listed on the BSE, and only four had recorded any trading during the preceding 365 days. Overall, 264 companies (94.3%) were delisted following the takeover. The median time to delisting was 355 days, with 50.4% leaving the market within one year, 62.7% within two years and 80.8% within five years. Delisting generally preceded legal exit: 187 of the delisted companies (70.8%) remained active, while the median period from takeover to deregistration among deregistered companies was nearly six years.

The more recent data show a similar pattern of ownership consolidation. Of the 54 target companies analysed for 2021–2026, 41 (75.9%) were subsequently subject to a squeeze-out and 42 (77.8%) were delisted, with squeeze-out followed by delisting in 39 cases. Pricing patterns are particularly relevant: in 58.1% of the 62 takeover bids, the shares did not meet the statutory liquidity requirement, making appraised fair value one of the statutory benchmarks for the minimum offer price. Of the 41 subsequent squeeze-outs, 34 (82.9%) were carried out at the takeover bid price. In 21 of these 34 cases (61.8%), market price had not served as the relevant statutory benchmark in the preceding takeover bid because the shares had failed the liquidity requirement, requiring an appraisal of fair value.

These findings confirm that persistent delisting remains a prominent feature of the BSE and show that post-takeover squeeze-outs are common in Serbia. Takeover pricing therefore frequently becomes the basis for subsequent compulsory minority exit, making the reliability and fairness of the underlying pricing process particularly important for minority shareholders.

3. Minority Shareholder Vulnerabilities and the Potential Role of Forensic Accounting: Insights from Selected Serbian Disputes

Minority shareholder disputes may ultimately require judicial or regulatory intervention where compensation, pricing or the acquisition of control is contested. Their resolution may depend on the reliability of financial information and projections, valuation inputs and expert opinions, as well as the transactions through which control, liquidity and the legally relevant price are established. These challenges are particularly relevant in Serbia, where conflicting valuations, limited judicial capacity to assess complex valuation issues and broader institutional constraints may weaken effective minority protection (Vasiljević & Radović, 2017; Radović, 2017; Mićović, 2024). The cases below illustrate these risks and the need for forensic accounting expertise.

The NLB Komercijalna banka case (NLB case; squeeze-out dispute over fair compensation; case R 95/2022 Commercial Court in Belgrade). The original valuation relied substantially on management-generated projections, including future dividend assumptions, which the appraiser accepted in full without independently testing their reasonableness or evidentiary basis. In subsequent proceedings, the court-appointed expert relied on a single valuation method (explicit dividend discount method) and determined the discount rate using a build-up approach uncommon in the valuation of banks, with bank-specific risk components set subjectively, without reference to comparable banks or companies.

A separate controversy concerned preferred shares. Public disclosures over a prolonged period, including stock-exchange prospectuses and takeover documentation, had consistently reported the same book value per share for ordinary and preferred shares, a treatment also publicly confirmed by the bank's Management and Executive Boards in their opinions on the takeover bid. For squeeze-out purposes, however, the preferred shares were treated as bond-like instruments with an effectively fixed return, although their terms provided only a minimum, rather than a fixed or maximum, dividend entitlement. On that basis, the appraiser, a reputable audit firm, treated their book value as equal to their nominal value, thereby reducing the book value per preferred share approximately 4.5-fold.

The Savremena administracija case (SVRM case; appraisal-rights dispute over fair compensation; case P 1588/2024 Commercial Court in Belgrade). One controversy concerned whether the application of the net asset method in share valuation could itself give rise to a new tax liability reducing the value of the shares. The court-appointed expert took the view that appraising real estate above its book value created such a liability. This reasoning, however, failed to distinguish valuation for the purpose of determining fair compensation from accounting treatment for financial reporting purposes (IAS 16). Deferred tax liabilities do not arise merely because shares are being valued; IAS 12 governs their recognition for financial reporting, and they do not even represent a current tax payment obligation,

but an accounting category. The court accepted the expert's interpretation, resulting in a substantial reduction in the compensation payable to dissenting shareholders.

The Energoprojekt Holding case (ENHL case; takeover and subsequent squeeze-out; case 26 U 2125/26(17) Administrative Court in Belgrade). One controversy concerned the ownership and trading relationships underlying the acquisition of control. In 2017, the Securities Commission found that formally separate investors had acted in concert in acquiring ENHL shares, based on ownership, management and funding links and coordinated trading activity, and required takeover bids for ENHL and two subsidiaries. The ENHL takeover was subsequently completed at RSD 1,501 per share, bringing the group's holding to 51.99%. The case also involved an insurance subsidiary in which indirect control was acquired without the required prior approval, which the National Bank of Serbia later refused to grant.

The pricing issue re-emerged in 2026, when a new takeover bid was launched at RSD 452.25 per share under the statutory rules applicable to liquid shares. The position of the Republic of Serbia, then the second-largest shareholder, added a further dimension. With the bidder's voting rights suspended, the Republic had sufficient voting power to determine profit allocation, yet voted to retain all undistributed profit. Shortly afterwards, it tendered its shareholding into the takeover bid, enabling the bidder to cross the 90% squeeze-out threshold. The contrast is notable: the state had declined the 2017 offer at RSD 1,501 per share but accepted RSD 452.25 in 2026 after voting to retain earnings corresponding to a potential dividend of approximately RSD 342.5 per share. Effectively, after accounting for the foregone dividend and based solely on direct cash flows in 2026, the state's net benefit from the transaction was approximately RSD 109.7 per share.

Almost nine years after the original takeover, and shortly before the 2026 bid closed, the Administrative Court annulled the Commission's 2017 decision approving the 2017 bid and returned the matter for reconsideration. By then, the original takeover had long been completed and the resulting ownership structure had underpinned subsequent corporate and regulatory developments. After the bidder crossed the 90% threshold in 2026 and initiated the squeeze-out, the Commission initially ordered that the transfer of shares should not proceed² (Komisija za hartije od vrednosti, 2026). Related judicial and regulatory proceedings remain ongoing.

Taken together, these cases suggest that minority shareholder vulnerability begins with the structural advantage of the controlling shareholder. Control may extend beyond voting rights to influence over management, dividend policy, related-party transactions, accounting choices and the financial information and projections later used in valuation. These factors may ultimately affect the price at which minority shareholders exit. A forensic review should therefore extend beyond the valuation

² By the time this paper was finalized, the squeeze-out had nevertheless been completed, although the related decision of the Securities Commission was not publicly available.

model to the information, assumptions and corporate decisions on which compensation ultimately rests.

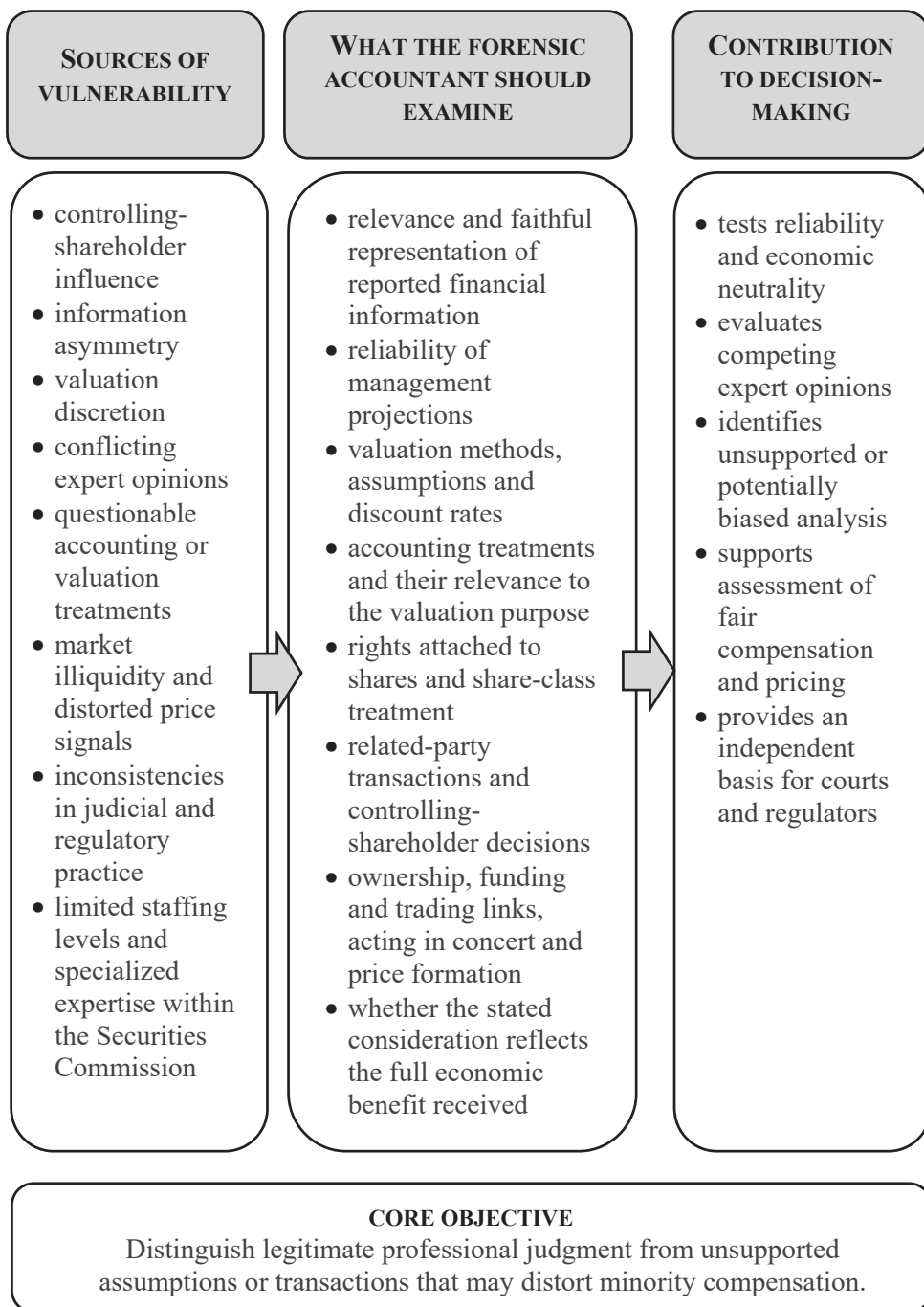
Picture 1: Selected Serbian Disputes and Minority Shareholder Risks

Case	Core controversy	What it illustrates
NLB Komercijalna banka (NLB)	Management projections; dividend-discount methodology and discount rate; treatment of preferred-share rights	Valuation inputs and expert discretion and expertise can materially affect compensation
Savremena administracija (SVRM)	Introduction of a tax liability arising from the valuation for fair compensation	Court appointment does not guarantee expert competence
Energoprojekt Holding (ENHL)	Acting in concert, ownership/funding links, price formation, state participation, inconsistencies in judicial and regulatory practice	The reliability of the takeover price may depend on transactions and relationships beyond the stated price, as well as on the conduct of public authorities, courts and regulators.

Source: Author's illustration

Professional and ethical standards require valuers and experts to act objectively. Even so, valuation standards leave considerable room for professional judgment in selecting methods, inputs, assumptions and adjustments. Where the initial valuer is engaged by the company or controlling shareholder, that discretion creates a structural risk of conscious or unconscious alignment with the engaging party's interests (Broekema et al., 2022). Judicial-appraisal research also links methodological choices and experts' professional backgrounds to valuation outcomes (Saastamoinen & Savolainen, 2019, 2021). The Serbian cases show that the problem can go further: expert discretion may extend beyond legitimate professional judgment to treatments difficult to reconcile with fundamental accounting or valuation principles. The SVRM and NLB cases illustrate this boundary. In SVRM, a financial-reporting recognition rule was applied to a valuation serving a different purpose. In NLB, the appraiser misinterpreted the nature and scope of the rights attached to preferred shares, while also departing from accepted valuation practice in applying the dividend discount method and determining the discount rate. In such circumstances, the forensic issue is no longer which of several professionally acceptable assumptions should prevail, but whether the expert treatment itself has a sound conceptual and evidentiary basis.

Minority shareholders may respond by commissioning their own expert and challenging the assumptions, methods or accounting treatments underlying the original valuation. The court may then face competing technical opinions and must determine not merely which figure is more persuasive, but which analysis rests on more reliable evidence and professional reasoning. This is particularly difficult where judges lack specialized valuation expertise (Radović, 2017). The difficulty is compounded where the court does not allow conflicting expert opinions to be tested against each other or the expert to be questioned (as in the SVRM case).

Figure 1: The Role of Forensic Accounting in Minority Shareholder Disputes

Source: Author's illustration

A court-appointed forensic accounting expert could therefore provide an independent layer of technical scrutiny rather than merely another valuation. The task would be to test the provenance and reliability of financial information and management projections, identify possible effects of related-party transactions or controlling-shareholder decisions, assess valuation methods, assumptions, and accounting treatments, and critically compare competing expert opinions. Court appointment alone, however, does not guarantee reliability: appropriate competence, professional skepticism and transparent, evidence-based reasoning remain essential.

The same logic applies to the Securities Commission. Where takeover pricing depends on appraised value, market liquidity or related-party transactions, forensic expertise could test whether the information and transactions underlying the statutory price are substantively reliable rather than merely formally compliant. This may include reconstructing ownership and funding links, coordinated trading and price formation. Forensic analysis should also examine whether the stated consideration represents the full economic benefit received by an accepting shareholder, or whether related transactions, side arrangements or other benefits alter the effective consideration. If such additional benefits exist, the stated takeover price cannot automatically be regarded as fair for shareholders who do not receive them (Kaisanlahti, 2007). Forensic accounting can thus provide courts and regulators with an independent basis for distinguishing legitimate professional judgment from unsupported assumptions or transactions that may distort minority compensation.

Conclusion

This paper suggests that, in Serbia's underdeveloped and illiquid market, formal compliance with statutory pricing rules alone may not ensure effective minority shareholder protection. The findings show that, in Serbia, takeovers are frequently followed by delisting and squeeze-outs, with the takeover price often serving as the basis for subsequent compulsory minority exit. In such circumstances, the reliability of the information, valuation inputs and transactions underlying that price becomes particularly important. The selected disputes further illustrate how controlling-shareholder influence, valuation discretion, conflicting expert opinions and questionable accounting or valuation treatments may materially affect minority compensation.

Forensic accounting can strengthen this protection by providing an independent examination that goes beyond producing another valuation. Its role may include testing management-generated information and projections, valuation methods and assumptions, accounting treatments, related-party transactions, ownership and funding links, trading patterns and whether the stated consideration reflects the full economic benefit received by accepting shareholders. This can assist courts and regulators in distinguishing legitimate professional judgment from unsupported assumptions or potentially biased analysis, and in identifying transactions that may distort minority compensation. However, court appointment alone does not ensure

reliability; competence, independence, professional skepticism and transparent, evidence-based reasoning remain essential.

The effectiveness of forensic accounting is also constrained by access to evidence. A privately engaged expert may identify inconsistencies but may lack access to internal accounting records, management projections, financing arrangements, communications or other non-public information necessary to test them. Court-appointed experts and regulators may therefore be better positioned where procedural or statutory powers allow access to such evidence. The contribution of forensic accounting ultimately depends not only on technical expertise, but also on the expert's independence, access to relevant evidence and procedural position. Strengthening forensic accounting capacity within judicial and regulatory processes, together with targeted forensic accounting training for judges and Securities Commission staff, could therefore improve the assessment of contested valuations and pricing outcomes and contribute to more effective minority shareholder protection in Serbia, with potential relevance to comparable emerging markets.

References

- Broekema, M. J. R., Strohmaier, N., Adriaanse, J. A. A., & van der Rest, J. P. I. (2022). Are Business Valuers Biased? A Psychological Perspective on the Causes of Valuation Disputes. *Journal of Behavioral Finance*, 23(1), 23–42. <https://doi.org/10.1080/15427560.2020.1821687>
- Denčić-Mihajlov, K. (2009). Reforms of Corporate Governance and Takeover Regulation: Evidence from Serbia. *South-Eastern Europe Journal of Economics*, 7(2), 205–227.
- Directive 2004/25/EC of the European Parliament and of the Council of 21 April 2004 on takeover bids. (2004). *Official Journal of the European Union*, L 142, 12–23.
- Kaisanlahti, T. (2007). When Is a Tender Price Fair in a Squeeze-Out? *European Business Organization Law Review*, 8(4), 497–519. <https://doi.org/10.1017/S1566752907004971>
- Komisija za hartije od vrednosti. (2026). *Rešenje br. 4/0-32-2561/52-17 od 05.08.2026*.
- Mićović, S. (2024). Zaštita prava manjinskih akcionara u postupku preuzimanja. *Poslovni izazovi*, 385–418.
- Mukwiri, J. (2013). Takeovers and incidental protection of minority shareholders. *European Company and Financial Law Review*, 10(3), 432–460.
- Mülbert, P. O., Sajnovits, A. (2025). Lowballing under the EU Takeover Bid Directive: Strategies, concerns, and gold-plating remedies. *European Business*

Organization Law Review, 26, 411–441. <https://doi.org/10.1007/s40804-024-00333-6>

Radović, V. (2017). Odmeravanje visine naknade nesaglasnim akcionarima. *Usklađivanje poslovnog prava Srbije sa pravom Evropske unije*, 64-90.

Saastamoinen, J., Savolainen, H. (2019). Does the choice in valuation method matter in the judicial appraisal of private firms? *Journal of Business Finance & Accounting*, 46(1-2), 183-199. DOI: 10.1111/jbfa.12364

Saastamoinen, J., Savolainen, H. (2021). Does a leopard change its spots? Auditors and lawyers as valuation experts for minority shareholders in the judicial appraisal of private firms. *Journal of Business Finance & Accounting*, 48(3-4), 613-636. DOI: 10.1111/jbfa.12436

Vasiljević, M., Radović, V. (2017). Teorijske osnove prava nesaglasnih akcionara. *Anali Pravnog fakulteta u Beogradu*, 65(3), 7-27.

Vlada Republike Srbije. (2021). *Strategija za razvoj tržišta kapitala za period od 2021. do 2026. godine* („Službeni glasnik RS“, br. 102/2021).

Zakon o preuzimanju akcionarskih društava. (2016). Službeni glasnik RS br. 46/2006, 107/2009, 99/2011 i 108/2016.

Zakon o privrednim društvima. (2025). Službeni glasnik RS br. 36/2011, 99/2011, 83/2014 - dr. zakon, 5/2015, 44/2018, 95/2018, 91/2019, 109/2021 i 19/2025.

Acknowledgements

The author gratefully acknowledges the anonymous reviewer for valuable comments and suggestions and financial support from the Ministry of Science, Technological Development and Innovation of the Republic of Serbia (Agreement No. 451-03-34/2026-03/200116).

Milan Čupić
Faculty of Economics and Business,
University of Kragujevac
ORCID: 0000-0002-0150-6682

Nikola Nikolić
Faculty of Economics and Business,
University of Kragujevac
ORCID: 0009-0009-9190-5887

Stefan Vržina
Faculty of Economics and Business,
University of Kragujevac
ORCID: 0000-0001-8372-591X

STOCK LIQUIDITY AND DIVIDEND PAYOUT OF JOINT- STOCK COMPANIES IN THE REPUBLIC OF SERBIA

UDK 336.76:336.64(497.11)
DOI <https://doi.org/10.54318/czf.2026.mc.019>

Abstract

This study examines the relationship between stock liquidity and dividend payout among joint-stock companies in the Republic of Serbia. The analysis is based on 176 firm-year observations for 25 companies traded on the Belgrade Stock Exchange between 2012 and 2025. Multiple panel regression models are applied to alternative measures of stock liquidity and illiquidity, as well as logit estimation and tests that account for the COVID-19-related economic shock. The results consistently indicate that stock liquidity is positively related to both the likelihood of dividend payment and the payout level. The findings suggest that stock liquidity and dividend payouts are complementary mechanisms in the Serbian emerging capital market, contributing to reduced information asymmetry and stronger investor confidence.

Key words: stock liquidity, dividend policy, dividend payout, emerging capital market

JEL Classification: G35, G12.

Introduction

Although dividend policy has attracted much research and retains its place in the economic literature, it is still not fully explained why firms pay dividends and why investors pay attention to them (Al-Hiyari, Kolsi, Lutfi, Shakkour, & Aljumah, 2024). This is known as the “dividend puzzle” (Black, 1976). Emerging capital markets, including the Serbian, add further pieces to the “dividend puzzle” and attract researchers seeking to explain listed companies dividend policies (Al-Hiyari et al., 2024; De Souza Junior, Hijazi, & Da Silva, 2025; Vržina & Čupić, 2026).

Given the complexity of dividend policy, the objective of the paper is to analyse the relationship between stock liquidity and the dividend policy of joint-stock companies in the Republic of Serbia. The research was conducted on a sample of 25 joint-stock companies whose shares were traded on the regulated markets of the Belgrade Stock Exchange (BSE) between 2012 and 2025. Two measures of stock liquidity were used: sign inverted log values of the Amihud illiquidity measure and trading volume. Two measures of dividend payout used in the paper are the dividend payout level and a dummy variable for the dividend payout decision.

The paper contributes to the existing literature in three ways. First, in the literature on dividends, determinants such as past dividend decisions, profitability, investment opportunities, free cash flow (Dewasiri et al., 2019), leverage, firm size, taxes (De Souza Junior et al., 2025), and corporate governance (Farooq, Al-Jabri, Khan, Ali Ansari, & Tariq, 2024) have been the focus of researchers. Stock liquidity as a determinant of dividend policy has been less examined (Banerjee, Gatchev, & Spindt, 2007; Jiang, Ma, & Shi, 2017; Stereńczak & Kubiak, 2022). Therefore, the paper contributes to expanding the literature on the relationship between stock liquidity and the dividend policy of joint-stock companies.

Second, as noted by Lai, Saffar, Zhu, and Liu (2020), Lobão et al. (2022) and Bai, Qin, and Bai (2024), the majority of studies on dividends are conducted on samples from developed capital markets, and many questions remain unanswered, especially with respect to emerging capital markets. The contribution of this paper is, therefore, the analysis of the relationship between stock liquidity and the dividend policy of joint-stock companies in the emerging capital market of Serbia. Generalisation of the results to other emerging capital markets may be justified, as they often share similar characteristics, such as higher volatility and less developed regulatory mechanisms (Čupić, Todorović, & Benković, 2023).

Third, the paper contributes to the literature on the information environment and information asymmetry, with implications for forensic accounting. Previous research suggests that stock liquidity can reduce information asymmetry by improving the information environment and strengthening external monitoring (Jiang et al., 2017; Hu et al., 2020). On the other hand, dividend policy may convey information to market participants, consistent with the dividend-signalling perspective (Ali Taher & Al-Shboul, 2023). This paper extends this literature by examining whether the relationship between stock liquidity and dividend policy documented in prior research is also evident in the Serbian capital market.

In addition to the introductory and concluding remarks, the paper consists of four parts. The first part provides a review of the literature on the determinants of joint-stock companies dividend policies. The second part presents the research methodology. The research results are presented in the third part. The fourth part presents a robustness analysis.

1. Theoretical background and literature review

In one of the first studies of the relationship between stock market liquidity and dividend policy, Banerjee et al. (2007) explain that investors in illiquid markets have a higher demand for cash dividends, while investors in highly liquid markets can cheaply create homemade dividends by selling their holdings in the firm. They therefore examine what they term the “liquidity hypothesis of dividend” assuming that stocks that pay cash dividends allow investors to satisfy their liquidity needs with less or no trading and thus enable them to avoid trading frictions. Banerjee et al. (2007) find strong empirical evidence to support this hypothesis on a sample of NYSE and AMEX firms for the period 1963 to 2003.

Bulan, Subramanian, and Tanlu (2007) find that NYSE, AMEX and NASDAQ firms are more likely to initiate a dividend when their stock illiquidity is higher, but that the change in illiquidity is not a significant predictor of initiations. They argue that the timing of dividend initiations is best explained by a synthesis of the maturity hypothesis with catering theory, given that they find that initiators are mature firms and that firms appear to time their initiations to coincide with periods when investor sentiment favours dividends. On the other hand, Kuo, Philip, and Zhang (2013) find that catering incentives persist only among firms in common law countries and not in civil law countries and that after adjusting for risk there is little support for catering theory even among firms incorporated in common law countries. A negative relationship between stock liquidity and dividend payouts is found in several more studies on firms from developed capital markets (Ali Taher & Al-Shboul, 2023; Kuo et al., 2013).

Lai et al. (2020) examine the relationship between stock liquidity and dividend payout in an international setting where countries have institutional frameworks that differ significantly from those in the US. Using a sample of firms from 52 countries, they find that the negative relationship between dividends and stock liquidity is stronger for firms in countries with sound political institutions and that the effect of political constraints on this relationship is more pronounced in countries with better legal institutions. They explain that a combination of sound political institutions and strong legal protection makes a difference to investors’ ability to extract dividends from low-liquidity firms and that when firms are facing financial constraints, investors have less bargaining power for more dividends when stock liquidity is low even in countries with strong political institutions.

From an agency theory perspective, on a sample of NYSE, AMEX, and NASDAQ firms, Hu, Huang, and Chen (2020) find that an increase in stock liquidity leads to higher dividend payouts and that the substitute relationship between low dividend payouts and weak creditor rights protection is the main cause. Lobão et al. (2022) explain that the results of studies focused on US and UK capital markets cannot be directly extrapolated to Continental European markets since there are important differences in corporate governance regimes that can influence firms’ financial choices. They examine the ability of trading volume to predict dividend changes and

find evidence of a positive relationship between unusually high volume and subsequent dividend changes in a sample of Spanish and Portuguese listed firms.

For listed firms within the Australian stock market implementing the tax imputation system, Nguyen (2020) and Bai et al. (2024) also find positive relationship. Bai et al. (2024) argue that most liquidity/dividend studies focus on countries with double tax systems such as the US, where both dividend income and capital gains are taxed at the investor level and where dividends are often subject to higher taxation compared with capital gains. This implies that investors prefer capital gains over cash dividends in periods of high market liquidity when trading costs are not excessively high. On the other hand, the imputation tax system allows shareholders to receive a tax credit for the full amount of the income tax paid by the firm, including any cash dividends which gives firms a stronger incentive to distribute cash dividends, as investors value the tax credit.

Results from several studies based on samples of firms from emerging economies are also conflicting. Jiang et al. (2017) hypothesise that stock liquidity lowers information asymmetry and increases insiders' incentive to pay dividends. Using a sample of Chinese listed firms during 2000–2014, they find that this relationship is more pronounced when the information environment is opaque, and when conflict between controlling shareholders and minority investors is severe. Similarly, using a sample of United Arab Emirates listed firms, Al-Hiyari et al. (2024) find that dividend payouts are negatively influenced by information asymmetry and that this relationship is less pronounced in firms with strong corporate governance. Nadia Sawitri and Sulistyowati (2018), however, find negative relationship, while Pourebrahimi and Seyed-Khosroshahi (2013) and Silva (2025) find no statistically significant relationship between stock liquidity and dividend payouts.

Using a sample of firms listed on the stock exchanges of fourteen Central and Eastern European economies, including Serbia, Stereńczak and Kubiak (2022) find a strong bidirectional relationship between liquidity and dividend policy. In other words, they find a strong positive relationship between stock liquidity and dividend policy, supporting the hypothesis that stock liquidity decreases information asymmetry between insiders and outsiders, thereby increasing incentives to pay dividends. At the same time, they find that stock liquidity is affected by a firm's dividend policy, indicating reverse causality between these two characteristics.

Based on the emerging market literature, a positive relationship between stock liquidity and dividend payouts is expected in the Serbian capital market. Studies from emerging economies often show that higher stock liquidity reduces information asymmetry and agency conflicts, thereby increasing firms' willingness to distribute cash dividends (Al-Hiyari et al., 2024; Jiang et al., 2017; Stereńczak & Kubiak, 2022). This expectation is particularly relevant for Serbia, where the capital market remains relatively small and illiquid, information asymmetry is pronounced, and only a limited number of listed firms distribute dividends regularly. Under such

market conditions, dividend payments may function as a mechanism for reducing uncertainty and strengthening investor confidence.

Serbian studies further suggest that dividend policy is characterised by persistence, with previous dividend payments representing a significant determinant of current payout decisions (Jovković, Vasić, & Bogičević, 2021), while institutional factors such as taxation also influence payout policies (Vržina & Čupić, 2026). Additionally, the value relevance of accounting information in Serbia implies that investors rely heavily on credible financial signals (Čupić et al., 2023), reinforcing the signalling role of dividends. Therefore, it is reasonable to hypothesise that firms with more liquid shares on the BSE will exhibit higher dividend payouts, even after controlling for firm-specific and macroeconomic variables.

2. Methodology

2.1. Sample

The sample structure is presented in Table 1. The empirical research included a total of 25 joint-stock companies that were actively operating in the Republic of Serbia in the period from 2012 to 2025.

Table 1: Sample

Year	Number of sample firms	Annual share turnover of sample firms (RSD)	Stock market capitalisation of sample firms (RSD)	Sample firms' share of BSE stock market capitalisation (%)
2012	14	5,498,801,741	198,408,460,594	30.90
2013	13	5,339,761,831	219,307,460,320	33.03
2014	14	5,741,825,871	234,708,109,449	32.85
2015	15	4,413,306,341	234,067,529,137	36.52
2016	16	3,185,403,630	217,815,757,132	36.77
2017	15	5,169,553,936	223,163,657,328	40.66
2018	15	4,378,448,537	227,539,854,394	43.53
2019	13	4,688,481,732	186,333,095,810	32.71
2020	6	811,984,947	55,069,470,550	10.52
2021	10	801,829,287	174,960,097,511	32.80
2022	12	2,088,316,791	228,579,602,975	55.81
2023	13	2,029,669,687	241,963,962,519	56.94
2024	12	2,406,824,847	266,992,839,569	55.54
2025	8	784,524,382	166,232,616,589	32.71

Notes: RSD is Republic of Serbia dinar

Only firms traded on the regulated markets of the Belgrade Stock Exchange (BSE), the only Serbian capital market, were included. The share of the market capitalisation of the sample firms in total BSE stock market capitalisation in every year of the analysed period is greater than 30%, with the exception of 2020, due to the impact

of COVID-19 and the decision to omit observations if there are fewer than 30 trading days for a given firm-year, following Banerjee et al. (2007) and Jiang et al. (2017). The sample includes a total of 176 observations.

2.2. Research model

To conduct the research, the dependent, independent, and control variables were identified and the corresponding regression models were defined. Similar to previous studies (Al-Hiyari et al., 2024; Hu et al., 2020; Jiang et al., 2017; Stereńczak & Kubiak, 2022), the analysis was conducted using multiple panel regression to determine the relationship between the defined independent variables and dividend payouts, using the following model:

$$DV_{it} = \beta_0 + \beta_1 DVOL_{it} + \beta_2 LIQ_{it} + \beta_3 SIZE_{it} + \beta_4 ROA_{it} + \beta_5 MBR_{it} + \beta_6 LEV_{it} + \epsilon_{it} \quad (1)$$

where DV represents the dependent variables. Following previous empirical studies (Banerjee et al., 2007; Jiang et al., 2017; Lai et al., 2020; Nguyen, 2020; Stereńczak & Kubiak, 2022), DP and DPL are defined as two dependent variables. DP is a dummy variable that takes the value 1 if the firm paid a dividend in the observed year and 0 otherwise. DPL represents the dividend payout level, measured as the ratio of the total amount of dividends paid to total assets.

Following previous studies (Banerjee et al., 2007; Jiang et al., 2017; Lai et al., 2020; Lobão et al., 2022; Nguyen, 2020), LIQ and $DVOL$ are defined as two independent variables related to stock liquidity. LIQ is based on Amihud's (2002) illiquidity measure interpreted as the daily price response associated with one monetary unit of trading volume and calculated as follows:

$$Amihud_{iy} = \frac{1}{D_{iy}} \sum \frac{|R_{iyd}|}{V_{iyd}} \quad (2)$$

where D_{iy} is the number of trading days for stock i in year y , while R_{iyd} is the stock return and V_{iyd} is the daily turnover in RSD (Republic of Serbia dinar) of stock i on day d in year y . Following Jiang et al. (2017) and Nguyen (2020), LIQ is determined by log-transforming $Amihud$, given that it is highly skewed, and multiplying it by -1 for ease of interpretation. In other words, $LIQ = -\ln(1 + Amihud_{iy})$. $DVOL$ is the logarithm of the market value of traded shares deflated by the Consumer Price Index (CPI), with 2006 as the base year:

$$DVOL_{iy} = \ln \ln \left[1 + \left(\sum_{t=1}^{D_{iy}} P_{iyd} \times N_{iyd} \right) \times \frac{CPI_{2006}}{CPI_y} \right] \quad (3)$$

where P_{iyd} is the share price and N_{iyd} is the number of common shares traded of firm i on trading day d in year y .

The control variables are *SIZE*, *ROA*, *MBR*, and *LEV*, measuring firm size, profitability, growth prospects, and indebtedness, respectively. *SIZE* is the natural logarithm of total assets, *ROA* is net income divided by total assets, *MBR* is the market-to-book ratio, and *LEV* is total liabilities divided by total assets. Following Jiang et al. (2017), these variables are included because large, profitable, less indebted and firms with few investment opportunities are more likely to pay dividends. A year dummy (*YEAR*) is included to remove aggregate time effects.

3. Research results

Table 2 presents the results of the descriptive statistical analysis. Dividends were paid in 120 (68.18%) of the 176 observations. The largest number of dividend-paying firms was recorded in 2016 (13 out of 16). The average log-transformed Amihud illiquidity measure is -0.5301 , while the average logarithmic trading volume is 12.87 . The average $\square ROA$ is positive and all firm-year observations report positive net income. The average *MBR* 1.38 , indicating that market values generally exceed book values. Finally, none of the firm-year observations has total liabilities exceeding total assets.

Table 2: Descriptive statistics

	Number of observations	Average value	Standard deviation	Minimum	Maximum
<i>DP</i>	176	0.6818	0.4671	0.0000	1.0000
<i>DPL</i>	176	0.0185	0.0217	0.0000	0.1453
<i>DVOL</i>	176	17.4237	1.8237	13.3938	21.1845
<i>LIQ</i>	176	-0.5301	0.6747	-2.7808	-0.0012
<i>SIZE</i>	176	22.9893	1.5924	19.7564	26.9921
<i>ROA</i>	176	0.0630	0.0618	0.0005	0.5772
<i>MBR</i>	176	1.3849	4.2757	0.0644	39.2255
<i>LEV</i>	176	0.2951	0.2404	0.0087	0.9138

Table 3: Multiple panel regression results

	DP		DPL	
C	-0.8098 (-0.6609)	-2.3062 (-1.9109)	0.1861 (0.9111)	0.0654 (0.3278)
LIQ	0.2675*** (5.4932)		0.0089*** (3.1065)	
DVOL		0.1077*** (4.9791)		0.0032*** (2.6335)
SIZE	0.0597 (1.0876)	0.0373 (0.6652)	-0.0073 (-0.8014)	-0.0045 (-0.4957)
ROA	0.9260** (2.1773)	0.7675 (1.7507)	0.0532** (2.1721)	0.0459 (1.8452)
MBR	-0.0143 (-1.1197)	-0.0083 (-0.6542)	0.0007 (0.3677)	0.0008 (0.4593)
LEV	0.2265 (0.8275)	-0.0221 (-0.0835)	-0.0011 (-0.0529)	-0.0206 (-1.0731)
YEAR	YES	YES	YES	YES
Adjusted R²	0.2292	0.2070	0.5330	0.5239
F-statistics	3.8903***	3.5378***	5.7551***	5.5852***
Method	RE	RE	FE	FE
Observations	176	176	176	176

Note: Coefficient estimates are reported first, with t-statistics in parentheses. RE is random-effects panel regression and FE is fixed-effects panel regression.

* Results are statistically significant at the 1% *** and 5% ** levels.

Table 3 presents the results of the multiple panel regression analysis. The Hausman test results indicate that the random-effects method should be applied for DP, while the fixed-effects method is appropriate for DPL. The results of the panel regression reveal that LIQ and DVOL are positively related to DP and DPL. Among the control variables, only ROA is significantly related to dividend payouts.

4. Robustness analysis

To verify the results, the research models were modified and additional tests were conducted. First, the variables related to stock liquidity were changed. Instead of *LIQ* and *DVOL* as liquidity variables, the illiquidity variables *ILLIQMA* and *NOTRD* were used. *ILLIQMA* is determined as follows:

$$ILLIQMA_{iy} = \frac{Amihud_{iy}}{AILLIQ_y} \#(4)$$

where $AILLIQ_y$ is the average value of $Amihud_{iy}$ for all stocks in the observed year. *NOTRD* is the ratio of the number of days without share trading to the total number of days in the observed year. The relationship between stock illiquidity and dividend payouts was tested using multiple panel regression analysis.

Table 4 presents the results of the multiple panel regression, which show that *ILLIQMA* is significantly and negatively related to *DP* and *DPL*, while *NOTRD* is significantly and negatively related only to *DP*. Among the control variables, only *ROA* is significantly and positively related to *DP* and *DPL*, confirming the results in Table 3.

Second, given that *DP* is a dummy variable that takes a value of 0 or 1, the relationship between stock liquidity and the dividend payout decision was also analysed using logit regression. Table 5 presents logit regression results showing that *LIQ* and *DVOL* are positively related to *DP*, confirming the panel regression results. Additionally, *MBR* is significantly and negatively related to *DP*.

Third, the year dummy was replaced with a COVID/post-COVID period dummy (CD), taking a value of 1 for 2020–2025 and 0 for 2012–2019. The relationship between stock liquidity and dividend payouts was analysed using multiple panel regression analysis. Table 6 presents multiple panel regression results showing that *LIQ* and *DVOL* are positively related to *DP* and *DPL*, confirming the panel regression results. Additionally, *ROA* is significantly and positively related to *DP* and *DPL*.

Table 4: Multiple panel regression results accounting for stock illiquidity measures

	DP		DPL	
C	-1.3407 (-1.1531)	0.1871 (0.1412)	0.1072 (0.5303)	0.1021 (0.4871)
ILLIQMA	-0.0935*** (-5.5651)		-0.0025*** (-2.5474)	
NOTRD		-0.7164*** (-4.2589)		-0.0145 (-1.5588)
SIZE	0.0795 (1.5253)	0.0228 (0.3899)	-0.0039 (-0.4332)	-0.0035 (-0.3691)
ROA	0.9339** (2.2058)	0.9517** (2.1658)	0.0518** (2.0911)	0.0520** (2.0688)
MBR	-0.0135 (-1.0930)	-0.0062 (-0.4838)	0.0010 (0.5166)	0.0018 (0.9858)
LEV	0.1897 (0.7138)	0.0348 (0.1270)	-0.0075 (-0.3598)	-0.0189 (-0.9230)
YEAR	YES	YES	YES	YES
Adjusted R²	0.2299	0.1748	0.5224	0.5081
F-statistics	3.9016***	3.0600***	5.5573***	5.3034***
Method	RE	RE	FE	FE
Observations	176	176	176	176

Note: Coefficient estimates are reported first, with t-statistics in parentheses. RE is random-effects panel regression and FE is fixed-effects panel regression.

* Results are statistically significant at the 1% *** and 5% ** levels.

Table 5: Logit regression results

	DP	
C	0.4799(0.1505)	-6.7230(-2.0712)
LIQ	1.3688*** (4.2521)	
DVOL		0.5803*** (3.8484)
SIZE	0.0329(0.2519)	-0.1486(-1.0206)
ROA	4.7538(1.1645)	4.7492(1.1645)
MBR	-0.4800** (-2.2015)	-0.5741** (-2.3883)
LEV	0.2850(0.3523)	0.8001(0.9542)
YEAR	YES	YES
McFadden R²	0.2009	0.1791
LR statistics	44.2377***	39.4223***
Observations	176	176

Note: Coefficient estimates are reported first, with z-statistics in parentheses.

* Results are statistically significant at the 1% *** and 5% ** levels.

Table 6: Multiple panel regression results accounting for COVID/post-COVID period

	DP		DPL	
C	-0.7836 (-0.7498)	-2.2519 (-2.1189)	0.1292 (0.8505)	0.0366 (0.7634)
LIQ	0.2594*** (5.4036)		0.0081*** (2.9170)	
DVOL		0.0947*** (4.5220)		0.0029*** (2.7862)
SIZE	0.0655 (1.4160)	0.0540 (1.1042)	-0.0047 (-0.7004)	-0.0027 (-1.2246)
ROA	0.8863** (2.1414)	0.7666 (1.7987)	0.0477** (2.1391)	0.0439** (2.0729)
MBR	-0.0135 (-1.2059)	-0.0061 (-0.5392)	0.0011 (0.6317)	-0.0001 (-0.2081)
LEV	0.0702 (0.2811)	-0.1713 (-0.6976)	-0.0072 (-0.3815)	-0.0201 (-1.7411)
CD	YES	YES	YES	YES
Adjusted R²	0.1812	0.1456	0.5446	0.0680
F-statistics	7.4567***	5.9697***	7.9771***	3.1269***
Method	RE	RE	FE	RE
Observations	176	176	176	176

Note: Coefficient estimates are reported first, with t-statistics in parentheses. RE is random-effects panel regression and FE is fixed-effects panel regression.

* Results are statistically significant at the 1% *** and 5% ** levels.

5. Discussion

The research results provide insights into the relationship between stock liquidity and dividend policy in the Serbian capital market. The positive stock liquidity-dividend relationship is in line with the evidence reported for China by Jiang et al. (2017) and for Central and Eastern European markets by Stereńczak and Kubiak (2022). It is particularly noteworthy that the latter study includes Serbia and documents a positive relationship between stock liquidity and dividend policy across Central and Eastern Europe.

The Serbian evidence does not support the liquidity hypothesis of dividends in its original form. Banerjee et al. (2007) argue that investors in illiquid stocks demand cash dividends because homemade dividends are costly, implying a negative relationship between stock liquidity and dividend payouts. Negative relationships have also been reported by Kuo et al. (2013) and Lai et al. (2020). The divergence of the results of the present research is not unexpected and confirms that the examined relationship is determined by legal protection, political institutions, taxation, corporate governance, and the information environment (Bai et al., 2024; Lai et al., 2020; Lobão et al., 2022). It would therefore be misleading to treat evidence from developed markets as a benchmark for Serbia.

The results are compatible with the explanation of Jiang et al. (2017) and Hu et al. (2020), which implies that stock liquidity can strengthen external monitoring and reduce information asymmetry, thereby increasing the incentives to distribute cash to shareholders. Dividend payments may provide an additional signal to outside investors regarding firms' financial condition (Ali Taher & Al-Shboul, 2023). These mechanisms may be particularly relevant in relatively small and less liquid capital markets, where information asymmetry can be more pronounced. The findings therefore suggest that stock liquidity and dividend policy may provide useful contextual information for forensic accountants when assessing the information environment and potential irregularities of listed companies.

The negative coefficient of *NOTRD* for *DP* suggests that stock illiquidity is more decisive for the dividend decision than for the dividend payout level. By contrast, *ILLIQMA*, which captures the price response to trading volume relative to the market, is significantly related to both dividend policy outcomes. No single proxy captures every dimension of stock liquidity, and alternative measures may capture different trading conditions (Jiang et al., 2017). The weaker *DPL* result for *NOTRD* should not be interpreted as evidence against the general relationship. Instead, it indicates that the decision whether to pay a dividend may be particularly sensitive to prolonged periods without trading.

The additional analyses support the robustness of the main result. *LIQ* and *DVOL* remain positive and statistically significant when the dividend decision is estimated by logit regression and when the models account for the economic shock. Profitability is the most consistent firm-level control, given that *ROA* is positive in

the baseline and alternative illiquidity specifications, although its significance varies across models. *MBR* is negative in the logit estimates but not in the panel regressions, while *SIZE* and *LEV* are generally insignificant. These patterns suggest that the stock liquidity coefficients are not merely capturing observable differences in firm size, profitability, growth opportunities, or leverage.

Conclusion

The objective of the research was to examine the relationship between stock liquidity and dividend policy in the Serbian capital market. It was conducted using 176 firm-year observations for 25 joint-stock companies traded on the regulated markets of the Belgrade Stock Exchange from 2012 to 2025. Across the baseline panel models, stock liquidity is related to both a higher probability of dividend payment and a higher dividend payout level. The alternative measures of stock illiquidity provide consistent evidence, showing that stock illiquidity is related to a lower probability of dividend payment, while the evidence for the payout level depends on the stock illiquidity measure used. The logit estimates and the models that account for the economic shock confirm the positive relationship between stock liquidity and dividend policy outcomes. The central conclusion is therefore that stock liquidity and dividend payouts are complementary in the Serbian capital market.

The research results have several practical implications. They indicate that managers of firms in emerging capital markets can reduce agency problems and information asymmetry, and consequently improve their reputation and access to finance, by enhancing stock liquidity and paying dividends. For investors, the findings suggest that stock liquidity should be considered when assessing potential market returns, as it can influence capital gains and dividend income. Forensic accountants can treat stock liquidity, trading activity and dividend decisions as contextual indicators when analysing the information environment and assessing potential irregularities. From the perspective of regulators and economic policy-makers, the results highlight the need to strengthen capital market institutions, primarily by improving transparency and the quality of corporate reporting and encouraging greater investor participation. The development of a more liquid and informationally efficient market can reinforce market signals in corporate decision-making and support more efficient resource allocation.

The research has certain limitations that also serve as guidelines for future studies. The analysis is based on a relatively small sample of firms, reflecting the limited number of liquid stocks on the Belgrade Stock Exchange. Although various indicators of stock trading and liquidity were used, it is possible that some aspects of market activity are not fully captured, leaving room for alternative measures and approaches. Future research could expand the analysis by including a larger number of firms from Serbia and other emerging markets. It would also be useful to examine the role of institutional investors, ownership structure and corporate governance in shaping dividend policy. Another venue for research may involve the application of

dynamic panel models, which would allow for a more detailed examination of the time dependence of dividend decisions. Future studies could also consider the interaction between dividend policy and other forms of profit distribution, such as share buybacks, which would contribute to a more comprehensive understanding of corporate payout policies in emerging markets.

References

- Al-Hiyari, A., Kolsi, M. C., Lutfi, A., Shakkour, A. S., & Aljumah, A. I. (2024). Information asymmetry and dividend payout in an emerging market: Does corporate governance quality matter? *Journal of Open Innovation: Technology, Market, and Complexity*, 10(1), 100188. <https://doi.org/10.1016/j.joitmc.2023.100188>
- Ali Taher, F. N., & Al-Shboul, M. (2023). Dividend policy, its asymmetric behavior and stock liquidity. *Journal of Economic Studies*, 50(3), 578–600. <https://doi.org/10.1108/JES-10-2021-0513>
- 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](https://doi.org/10.1016/S1386-4181(01)00024-6)
- Bai, M., Qin, Y., & Bai, F. (2024). Dividend policy and firm liquidity under the tax imputation system in Australia. *International Journal of Managerial Finance*, 20(2), 281–303. <https://doi.org/10.1108/IJMF-01-2023-0018>
- Banerjee, S., Gatchev, V. A., & Spindt, P. A. (2007). Stock Market Liquidity and Firm Dividend Policy. *Journal of Financial and Quantitative Analysis*, 42(2), 369–397. <https://doi.org/10.1017/S0022109000003318>
- Black, F. (1976). The Dividend Puzzle. *The Journal of Portfolio Management*, 2(2), 5–8. <https://doi.org/10.3905/jpm.1976.408558>
- Bulan, L., Subramanian, N., & Tanlu, L. (2007). On the Timing of Dividend Initiations. *Financial Management*, 36(4), 31–65.
- Čupić, M., Todorović, M., & Benković, S. (2023). Value relevance of accounting earnings and cash flows in a transition economy: The case of Serbia. *Journal of Accounting in Emerging Economies*, 13(3), 541–565. <https://doi.org/10.1108/JAEE-12-2021-0411>
- De Souza Junior, W. D., Hijazi, M. M., & Da Silva, T. P. (2025). Determinants of Dividend Payout Policy: More Evidence From Emerging Markets of G20 Bloc. *International Journal of Finance & Economics*, 30(4), 4113–4124. <https://doi.org/10.1002/ijfe.3111>

- Dewasiri, N. J., Yatiwelle Koralalage, W. B., Abdul Azeez, A., Jayarathne, P. G. S. A., Kuruppuarachchi, D., & Weerasinghe, V. A. (2019). Determinants of dividend policy: Evidence from an emerging and developing market. *Managerial Finance*, 45(3), 413–429. <https://doi.org/10.1108/MF-09-2017-0331>
- Farooq, M., Al-Jabri, Q., Khan, M. T., Ali Ansari, M. A., & Tariq, R. B. (2024). The impact of corporate governance and firm-specific characteristics on dividend policy: An emerging market case. *Asia-Pacific Journal of Business Administration*, 16(3), 504–529. <https://doi.org/10.1108/APJBA-01-2022-0007>
- Hu, Y., Huang, W., & Chen, Z. (2020). Stock liquidity, agency cost, and dividend payouts. *Applied Economics Letters*, 27(4), 335–339. <https://doi.org/10.1080/13504851.2019.1616052>
- Jiang, F., Ma, Y., & Shi, B. (2017). Stock liquidity and dividend payouts. *Journal of Corporate Finance*, 42, 295–314. <https://doi.org/10.1016/j.jcorpfin.2016.12.005>
- Jovković, B., Vasić, A. S., & Bogićević, J. (2021). Determinants of Dividend Policy: A Case of Serbia's Banking Sector. *Naše gospodarstvo/Our Economy*, 67(1), 13–22. <https://doi.org/10.2478/ngoe-2021-0002>
- Kuo, J.-M., Philip, D., & Zhang, Q. (2013). What drives the disappearing dividends phenomenon? *Journal of Banking & Finance*, 37(9), 3499–3514. <https://doi.org/10.1016/j.jbankfin.2013.05.003>
- Lai, K. M. Y., Saffar, W., Zhu, X. (Kevin), & Liu, Y. (2020). Political institutions, stock market liquidity and firm dividend policy: Some international evidence. *Journal of Contemporary Accounting & Economics*, 16(1), 100180. <https://doi.org/10.1016/j.jcae.2019.100180>
- Lobão, J., Piedade, P., & Nippani, S. (2022). Does stock trading volume signal future dividends? Evidence from Iberian firms. *Portuguese Economic Journal*, 21(1), 53–66. <https://doi.org/10.1007/s10258-020-00191-3>
- Nadia Sawitri, A., & Sulistyowati, C. (2018). Stock Liquidity and Dividend Policy. *KnE Social Sciences*, 3(10). <https://doi.org/10.18502/kss.v3i10.3461>
- Nguyen, T.-G. (2020). Stock liquidity and dividend policy: Evidence from an imputation tax environment. *International Review of Financial Analysis*, 72, 101559. <https://doi.org/10.1016/j.irfa.2020.101559>
- Pourebrazhimi, M. R., & Seyed-Khosroshahi, S. A. (2013). The Relationship between Dividend Policy and Trading Volume in Tehran Stock Exchange. *Financial Research Journal*, 14(2), 15–30. <https://doi.org/10.22059/jfr.2013.50626>
- Silva, A. R. D. (2025). Do trading volume and volatility influence dividend distributions? A study of companies in the electric sector in brazil and the usa. *FACEF Pesquisa - Desenvolvimento e Gestão*, 28(3). <https://doi.org/10.29327/2681802.28.3-5>

Stereńczak, S., & Kubiak, J. (2022). Dividend policy and stock liquidity: Lessons from Central and Eastern Europe. *Research in International Business and Finance*, 62, 101727. <https://doi.org/10.1016/j.ribaf.2022.101727>

Vržina, S., & Čupić, M. (2026). Poreski aspekti raspodele dobitka preduzeća u Srbiji. *Zbornik Radova*, 319–333. Ekonomski fakultet Univerziteta u Kragujevcu. <https://doi.org/10.46793/RZ26.319V>

Snežana Miletić
University of Belgrade, Faculty of Economics
and Business
ORCID: 0000-0001-9116-9386

Siniša Jović
Serbian Tax Administration, Tax Police
Department

ELEMENTS OF MTIC AND CAROUSEL VAT FRAUD OUTSIDE THE EU: A FORENSIC COMPARISON

UDK 336.2:343.72
DOI <https://doi.org/10.54318/czf.2026.sm.020>

Abstract

This paper examines which elements of missing trader intra-Community (MTIC) and carousel-type value-added tax (VAT) fraud, as documented in European Union practice, can be identified in VAT refund fraud detected in a non-EU jurisdiction.

Two forensic case studies from Serbia are compared against a predefined set of features, separating elements common to both contexts from those dependent on EU law. The first case involves an export scheme in which goods were purchased domestically at artificially inflated prices and exported to a third country, generating refund claims of approximately 48 million Serbian dinars (RSD), of which approximately RSD 32 million was refunded, with a zero-rated export supply but no missing trader. The second involves a closed network of related domestic entities in which the same goods circulated repeatedly over several years, generating cumulative VAT refunds of approximately RSD 500 million without any cross-border transaction, while simulated retail sales lent credibility to inflated prices.

The comparative analysis identifies three conditions present in both cases and in the EU model: a systematic excess of input VAT over output VAT, artificial inflation of transaction values within a coordinated network, and the absence of a corresponding cash VAT payment in the period concerned. Elements dependent on the EU legal framework, such as intra-Community acquisition and jurisdictional fragmentation, were absent from both cases, as was the disappearance of a trader. In the first case suppliers offset output VAT against carried-forward import VAT credits. In the second, participants reduced their reported net VAT liabilities by claiming fictitious input VAT from invoice mills. The paper's contribution is to show that exposure to this type of fraud follows from the invoice-credit mechanism itself rather than from EU law, and to propose a typology based on the mechanisms used to obtain unwarranted VAT refunds rather than on EU membership.

Key words: MTIC fraud, carousel fraud, VAT refund fraud, invoice-credit mechanism, invoice mills, Serbia, case study.

JEL Classification: H26, K42, M49

Introduction

Value added tax has become an almost universal form of consumption taxation. By 2023 it had been adopted in 175 jurisdictions, compared with approximately 30 in the early 1980s, the United States remaining the only major economy not to have introduced it (Brockmeyer et al., 2024, p. 107). Part of the explanation for this rapid spread lies in the belief that the invoice-credit mechanism makes the tax largely self-enforcing. Every participant in the chain has an interest in obtaining a proper invoice from its supplier, and the resulting documentary trail assists the tax administration in verification. That same mechanism, however, opens a form of abuse unknown to a single-stage retail sales tax.

Keen and Smith (2006, pp. 866-867) distinguish between forms of fraud that arise equally under a VAT and under a single-stage retail sales tax, such as the under-reporting of sales, the failure to register, or the failure to remit tax collected, and forms specific to the credit mechanism, in which the taxable person claims a deduction or refund of input VAT without entitlement, on the basis of fictitious documentation or overstated invoices. In describing the second category the authors invoke Bird, for whom a VAT invoice is "a check written on the government" (Bird, 1993, as cited in Keen & Smith, 2006, p. 867). Tait (1988, p. 307) records an estimate that in the Netherlands 44% of all VAT fraud took the form of fake claims for prior-stage tax, most often through fabricated invoices for purchases never made.

The assumption of self-enforcement holds only in part. The purchaser has an interest in the reality of the transaction, on which its own right of deduction depends, but not in whether the supplier has actually remitted the amount stated, a fact the purchaser cannot ordinarily verify (Keen & Smith, 2006, p. 865). Empirical work confirms that the paper trail has a deterrent effect, but that its reach is limited. Drawing on a sample of over 445,000 Chilean firms, Pomeranz (2015, p. 2540) shows that the announcement of increased scrutiny has a weaker effect on transactions that already leave a documentary trail, which points to the preventive deterrent effect of that trail. Naritomi (2019, p. 3033), examining a consumer reward programme in São Paulo, finds that reported revenue in retail sectors rose by at least 21% over four years, that is, only once the consumer had been given an incentive of its own to demand a receipt. The gap between the formal correctness of an invoice and the supplier's actual discharge of the liability is what makes fraud of this kind possible, and it is also why detection cannot rest on documentary verification at the level of the individual taxable person.

The most serious form of abuse of the credit mechanism is described in the literature as missing trader intra-Community fraud, and in its more elaborate form as carousel fraud. Both, however, are almost invariably examined using evidence from the European Union, modelled on trade between Member States and analysed through the case law of the Court of Justice, and some authors go so far as to assert that the phenomenon does not occur outside that framework.

The research question is therefore the following: which elements of the MTIC and carousel model are conditioned by the legal architecture of the European Union, and which follow from the invoice-credit mechanism itself and may therefore arise where that architecture is absent? What the paper examines is opportunity in the sense of the fraud triangle, that is, one of the three conditions for fraud set out by Cressey (1953), namely the structural conditions that make the commission of these offences possible, traced through the schemes analysed and compared across differing legal jurisdictions.

The paper answers this question through a forensic analysis of two case studies drawn from files of the Tax Police of the Ministry of Finance of the Republic of Serbia. The cases were selected purposively, to provide a contrast. The first contains a cross-border element and a zero-rated supply but no participant that disappears. The second contains no foreign transaction at all, yet does contain participants performing the function of a missing trader and the actual circulation of the same goods. Each case was reconstructed element by element from VAT returns and records, accounting and customs documentation, field inspection reports, expert findings, data on cross-border payments and statements obtained in the proceedings. The identity of the participants is protected and the entities are therefore designated by codes. The material was anonymised before use and no participant can be identified from it, so that no specific authorisation was required for its use in research¹. The unit of analysis is the individual scheme, and every element was verified by cross-checking at least two independent sources. The reconstructed schemes were then compared with a model derived from EU literature and legislation, against a predetermined set of attributes, so that for each attribute it could be established whether it depends on the EU legal framework. The theoretical part of the paper draws on the relevant academic literature, EU legislation, and the case law of the Court of Justice of the European Union and of the European Court of Human Rights.

A forensic approach here means that the conclusions are derived by cross-checking tax records against inspection findings, expert examination, witness statements and financial flows. The distinction matters, because an analysis of returns alone reveals how much revenue was lost but not how the steps of the scheme are connected to one another. The literature at the intersection of taxation and forensic accounting remains sparse, even though it is precisely there that questions of reconstructing taxable income and proving tax fraud arise (DiGabriele et al., 2024, pp. 213-215), and the engagement of forensic accountants in tax disputes is documented in the case law as well (Muehlmann, Burnaby & Howe, 2012, pp. 1-3).

¹ The two cases are drawn from files in which the pre-investigation measures resulted in the deprivation of liberty of nine and of thirty-three persons respectively, on complaints alleging VAT fraud, money laundering and, in the second case, criminal association. The outcome of those proceedings is outside the scope of this paper.

1. Theoretical framework

1.1 Missing trader intra-Community (MTIC) fraud: concept and mechanism

Missing trader intra-Community (MTIC) fraud emerged, and has been described, in the context of the EU single market, in the rules governing the taxation of trade in goods between Member States and in the literature concerned with that taxation. These sources do not define the concept uniformly. What follows therefore considers both the general and the legal definition, and then the mechanism by which the fraud arises.

In general terms, a missing trader (MT) is a taxable person registered for VAT purposes who, having stated VAT on an output invoice, fails to remit that amount to the tax authority and becomes untraceable (Podlipnik, 2012, p. 463). EU legislation defined the concept somewhat more narrowly. Under Article 2 of Commission Regulation (EC) No 1925/2004, which ceased to apply in 2012, a missing trader is a taxable person registered for VAT who, potentially with fraudulent intent, acquires or purports to acquire goods or services without paying VAT and then supplies those goods or services with VAT stated on the output invoice, without remitting the amount due to the competent national authority. That definition differs from the general one in its criteria: it requires the acquisition to have been made without payment of VAT, but it does not require the taxable person to become untraceable. Regulation (EU) No 79/2012, which replaced it, contains no definition of the missing trader, so that no binding legal definition of the concept exists today, although it is defined operationally in documents of the European Commission and Europol.

The basic mechanism of MTIC fraud follows directly from the way EU law governs taxable trade in goods between Member States. Where a taxable person in one Member State supplies goods to a taxable person registered for VAT in another, that supply is, subject to the prescribed conditions, exempt from VAT on the part of the supplier, who states no VAT on the output invoice (Article 138 of Directive 2006/112/EC). The taxable person in the Member State of destination must account for VAT on that intra-Community acquisition² as its own liability (Article 200), while at the same time, in the same return, becoming entitled to state the same amount as input VAT (Article 168(c)). This mechanism is known in the legislation and the literature as the reverse charge. The liability and the right of deduction thus cancel one another, and no actual payment of VAT to the revenue authority arises on that account. Because the two amounts offset each other, the acquirer is left with no net input VAT against which to reduce the liability on its next sale, so that the amount it retains if it disappears before payment corresponds to the whole of the VAT charged on that sale.

² Directive 2006/112/EC continues to use the expression intra-Community acquisition of goods, although since the entry into force of the Treaty of Lisbon the organisation has been named the European Union. In this paper the term Community is used only when that legal institution is named; elsewhere the term European Union, or EU, is used.

The distinction at issue is not one between Serbia and the EU as systems, but between two regimes: intra-Community acquisition, and importation from a third country, on which VAT is as a rule actually paid within the EU as well. In the Republic of Serbia, the Law on Value Added Tax exempts exports of goods from VAT (Article 24(1)(2)) while preserving the right to deduct input VAT (Article 28(1)(2)). On importation, however, VAT must actually be assessed and paid, since the person liable is the importer (Article 10(1)(5)), and the acquirer becomes entitled to deduct only upon holding proof that the VAT stated has been paid (Article 28(2)(2)). Unlike intra-EU trade, where the tax charged and the tax deductible cancel out in the same return without any payment whatever, here payment precedes deduction. This is not a technical detail. It explains why, in intra-EU trade, input VAT can arise that was never preceded by a payment into the budget, whereas in the Serbian system an import credit arises only upon actual payment. The comparison concerns the regime that applied in the cases analysed. Article 211 of Directive 2006/112/EC permits Member States to provide that import VAT need not be paid at the time of importation but may instead be entered in the VAT return, so that in those States deduction is not preceded by payment on importation either. The difference therefore follows from a national choice of implementation rather than from EU membership, which further supports the paper's central claim.

In its simplest form, MTIC fraud involves a single passage of the goods through the chain: the goods enter the system once, the perpetrator obtains an unlawful gain, and the chain ends there. In practice this basic model is frequently elaborated, in that the goods, or at least the documentation relating to them, are not sold in the market of the Member State of destination but are re-exported and returned to the beginning of the chain, the whole process being repeated a number of times. This form is known as carousel fraud.

1.2 Carousel fraud as a complex form of MTIC

In carousel fraud the chain does not break once the missing trader has obtained its unlawful gain and disappeared. Instead, through a further series of transactions, the goods are returned to the undertaking whose supply set the chain in motion, and the circle closes. The whole process can be repeated with the same goods, so that each new round produces a fresh unlawful gain.

The classic carousel model involves four types of participant: the undertaking that sets the circle in motion by making a zero-rated supply of goods to another Member State; the missing trader, which acquires the goods without paying VAT to its supplier and disappears before remitting the VAT it has collected; one or more buffer traders through which the goods pass with VAT duly stated and remitted; and the undertaking that, at the end of the chain, supplies the goods back and claims a refund of input VAT. The structure of this model, set out in Table 1, is derived from the numerical example given by Keen and Smith (2006, pp. 871-872).

Table 1: Numerical example of the classic carousel model (amounts in EUR; purchases and sales exclusive of VAT)

Role	Member State	Purchase price	Sale price	Output VAT	Tax treatment	
					VAT not remitted by B	VAT remitted by B
Undertaking A (supply to B)	X	—	10,000	0	Exempt with deduction	Exempt with deduction
Undertaking B (MT)	Y	10,000	11,000	2,200	Does not remit	Remits
Undertaking C (buffer)	Y	11,000	12,100	2,420	Remits 220	Remits 220
Undertaking D (supply to A)	Y	12,100	13,310	0	Refund 2,420	Refund 2,420
Effect on the budget of State Y					–2,200	0

Source: authors' numerical example, constructed on the logic of the examples in Podlipnik (2012, pp. 463-464) and Keen and Smith (2006, pp. 871-872).

To illustrate the classic carousel model, assume that each undertaking in the chain applies a mark-up of 10% to its purchase price and that taxable domestic supplies are subject to a VAT rate of 20%, as set out in Table 1.

The table shows the flows and the amounts, but not the reason why the whole of the VAT that undertaking B charges to undertaking C constitutes its net liability. Undertaking B paid no VAT to any supplier, and the amount it accounted for on the acquisition has already cancelled its own liability on that account, so that no deduction remains available on the onward sale. If it disappears before discharging that liability, it retains it in full, and the revenue loss equals that amount.

The mark-ups applied by undertakings C and D along the chain do not affect the size of the loss, because the tax that undertaking C remits on its own mark-up is cancelled by the identical amount that undertaking D claims on the same account. The example uses ordinary mark-ups in order to isolate the effect of the trader's disappearance. In practice, prices within a carousel are deliberately raised, since the missing trader fails to remit tax calculated on precisely that inflated price, so that the amount retained is proportionate to it. The goods must then periodically be undervalued if the circle is to be repeated (Financial Action Task Force [FATF], 2007, pp. 3-4). The structure of four roles and a closed circle is therefore no more than the framework within

which the loss may arise, while its actual cause is that at one point in the chain tax was collected from the purchaser and not remitted to the budget. Whether that outcome can be reached by means other than the disappearance of a trader is shown by the case studies from the Republic of Serbia.

Fedeli and Forte (2012, pp. 58-61) model the carousel as a chain of bilateral negotiations between each pair of participants, in which the price and the division of the gain obtained through unpaid VAT are agreed, each negotiation being resolved as a Nash bargaining solution. Across all nine numerical scenarios they test (2012, p. 67, Table 3), the greater part of total evasion, between approximately 96.88 and 98.38%, is realised at the final stage of the chain, in the State in which the refund is claimed, while the share of evasion in the State of origin is negligible. That concentration follows from the structure of the repeated tax credit rather than from any requirement that the final stage be a cross-border transaction, which raises the question whether the trigger for the refund must be an export or whether another event with the same fiscal effect can take its place³.

Further confirmation that the essential element of the carousel is tied neither to a particular type of good nor to physical cross-border movement is provided by the spread of this form of fraud to services and intangibles. Borselli, Fedeli and Giuriato (2015, pp. 710-711) describe the first known VAT fraud in the electronic services sector, carried out in Italy, in which the subject of the carousel was rights to digital content and telecommunications services, while Jakobsson (2013, pp. 19-24) describes a carousel within the EU Emissions Trading Scheme, in which carbon allowances circulated. Both examples show that the model can be applied to goods, services or intangible rights capable of generating a right to deduct or to a refund of VAT, which shifts the weight of the explanation from the nature of the goods to the capacity of the transaction to produce a tax credit.

Two elements emerge from the foregoing as the working assumption that the case studies are designed to test. The first is a mechanism that, at some stage of the chain, relieves the supply of VAT and thereby opens the possibility of a refund. The second is the capacity of the participants to inflate artificially, within a closed circle, the values on which the refund is calculated. What remains to be established is whether

³ An ordinary domestic sale would not suffice here, because the output and input VAT of undertaking D would largely cancel out without any actual cash refund, so that the loss would remain limited to the amount already retained by the missing trader. What is required is an event that, like an export, gives undertaking D a right to an actual refund. Such an effect may be produced, for example, by the false presentation of a domestic sale as an export, known in the literature as diversion fraud (Keen & Smith, 2006, p. 868); by the artificial inflation of the value of purchase invoices relative to sales invoices within a closed circle, which generates a permanent excess of deductible tax; or by the application of a special tax treatment, such as a zero rate or a supply exempt with the right to deduct, to a particular category of domestic trade.

these two elements are necessarily tied to the legal architecture of the EU, and whether both are indispensable.

1.3 The EU legal architecture as a precondition of the classic model

Exemption from VAT is not in itself unusual, since every system that taxes exports on the destination principle provides for it. What makes intra-EU trade particularly exposed is the place at which the chain is broken. Keen (2007, p. 17) locates the problem in the break in the VAT chain that occurs when goods pass between States. That is a particularly unfortunate place for the chain to break, because it is precisely there that control passes from one national tax administration to another. The break therefore occurs not at the boundary between two stages of trade but at the boundary between two jurisdictions, neither of which has a complete view of the transaction chain.

That state of affairs is a consequence of the transitional arrangement of 1993, when Council Directive 91/680/EEC abolished fiscal frontiers between Member States on the understanding that the transitional regime would later be replaced by a definitive system (Article 402 of Directive 2006/112/EC). In the absence of customs formalities at internal frontiers, verification rests on documentation and on the exchange of information between tax administrations, for which the VAT Information Exchange System (VIES) was established. Its use was, on 2005 figures, disproportionate to the volume of trade: some 26,000 exchanges of information took place, although roughly 35 million taxable persons were engaged in intra-Community trade (House of Lords European Union Committee, 2007, p. 18). This describes the state of the system at that time and is not an assessment of its present capacity.

A second feature of this architecture is the way EU law allocates risk between the State and the taxable person. In the joined cases *Optigen and Others* (C-354/03 and others, judgment of 12 January 2006, paragraph 52), the Court of Justice held that a taxable person who neither knew nor could have known of fraud elsewhere in the chain retains the right of deduction, whereas in *Kittel and Recolta Recycling* (C-439/04 and C-440/04, judgment of 6 July 2006, paragraph 61) that right was denied to a person who knew or ought to have known that it was participating in a transaction connected with fraud. Article 205 of Directive 2006/112/EC further permits Member States to impose joint and several liability on another person, and a limit to such measures was set by the European Court of Human Rights in *Bulves AD v Bulgaria* (application no. 3991/03, judgment of 22 January 2009, paragraph 71), which held that denying deduction to a taxable person acting in good faith on account of its supplier's default infringes the right to property.

All of the above explains why proposals for solving the problem within the EU are as a rule structural in nature. A general reverse charge mechanism, the consistent application of the destination principle with a single point of declaration, and taxation at origin with a clearing house have all been considered, each such change requiring unanimity among the Member States (House of Lords European Union

Committee, 2007, pp. 24-29). None had been adopted by the date of that report, because each removes one weakness by introducing another. The first would shift the entire tax liability onto the final link before final consumption, thereby depriving VAT of the very property for which it was introduced (Keen, 2007, p. 16); the second would impose unacceptable compliance costs on business; and the third would leave every State dependent on collection carried out by another administration.

Empirical assessments of the limited measures that have been implemented point in the same direction. Using German data, Buettner and Tassi (2023, pp. 849, 857) estimate revenue losses from fraud in the sectors to which the reverse charge was subsequently extended at about 5% of VAT revenue. Bussy and Tassi (2025, p. 1) carry out the same comparison on a sample of several Member States. More significant for present purposes is their finding that fraud is concentrated in States with higher levels of corruption, weaker customs efficiency and lower gross domestic product per capita. Within the sample of Member States observed, the extent of fraud therefore tracks the quality of the tax and customs administration. No conclusion about the significance of membership itself follows from that finding, since membership does not vary within the sample.

What emerges is that the absence of customs control at the internal frontier, the reliance on the exchange of information between two administrations of equal standing, and the rules on the liability of a taxable person acting in good faith developed in the case law of the Court of Justice are all conditioned by the legal architecture of the EU. The question is whether the elements without which the scheme cannot function are conditioned by it as well.

1.4 The representation of non-EU cases in the literature on MTIC and carousel fraud

MTIC and carousel fraud are almost invariably examined using evidence from the EU, and it is on that material that they are modelled (Fedeli & Forte, 2012, pp. 55-70) and measured (Buettner & Tassi, 2023, p. 849), while the legal analysis proceeds through the case law of the Court of Justice of the European Union. Some authors explicitly argue that the phenomenon does not occur outside that framework. Podlipnik (2012, p. 458) states that MTIC arises because of a fundamental flaw in the design of the EU VAT system, and that this type of fraud is not known in other, non-EU, tax systems that tax consumption⁴.

⁴ In a footnote to that same passage, Podlipnik (2012, p. 458, n. 7) states that, generally speaking, every VAT system is susceptible to missing trader fraud, since in every system there are circumstances in which trade between businesses is carried out without VAT being charged. The two statements operate at different levels of the concept: the first concerns MTIC as a form tied to intra-Community acquisition, the second missing trader fraud in general. This reading is supported by the legal definition in Article 2 of Commission Regulation (EC) No 1925/2004, which also covered a taxable person that merely purports to acquire the goods, and thus a wider category than the model presupposes.

That the missing trader phenomenon is not confined to the EU is confirmed by the findings of other authors and of international bodies. Tait (1988, pp. 307-308) describes both mechanisms five years before fiscal frontiers between Member States were abolished. He records that companies were established solely in order to issue false invoices to those wishing to defraud the revenue, and, separately, the newly established business that claims large refunds and then ceases to trade and becomes untraceable before its turn for audit arrives. Neither description presupposes cross-border trade. Keen and Smith (2006, p. 862) observe that VAT fraud and evasion are not confined to EU Member States or to high-income countries, citing Sri Lanka, which lost about 10% of its net VAT receipts through a single fraudulent episode. The same emerges from the FATF report, which took the EU as its subject but found that the same cross-border schemes exist in other States as well, expressly naming Mexico and Ukraine (FATF, 2007, p. 2).

The phenomenon has moreover been examined outside the EU, but by different methods, namely the analysis of tax administration data and the evaluation of reforms by difference-in-differences estimation, and for a different purpose. Using VAT return data from Pakistan, Shah (2026, pp. 1, 3) finds that approximately 10% of the firms covered by the reform were fake, created for the purpose of missing trader fraud, and notes that in developing countries, which charge tax on every import without exception as to the origin of the goods, domestic variants of that fraud may exist, which he terms domestic missing trader fraud. Examining the same country, Waseem (2023, p. 1) finds that VAT evasion in the sectors covered amounts to between 31 and 46% of potential revenue, that overclaimed refunds account for between 11 and 23% of that revenue, and that roughly two-fifths of the refunds so claimed rest on spurious invoices issued by companies established solely in order to issue them.

The question is not a peripheral one, given that VAT is now an almost universal form of consumption taxation, and that fraud in cross-border trade within the EU is concentrated in States of weaker institutional capacity, which points to the significance of the institutional capacity of the tax and customs administration.

What the literature reviewed lacks is therefore not research on MTIC and carousel fraud outside the EU, but forensic reconstruction of individual schemes element by element. The existing non-EU studies work with administrative data and measure the effects of reforms, so that it cannot be seen from them which specific elements of the model appear in those cases and which do not. For the Republic of Serbia there exists a forensic reconstruction of a VAT refund fraud scheme (Miletić & Popović, 2026, pp. 141-143), but it is concerned with the abuse of corporate vehicles and the connection with money laundering rather than with comparison against the MTIC and carousel model, so that the question of which elements of that model appear in domestic schemes, and which are absent, remains open.

2. Unwarranted VAT refunds without a missing trader: an export scheme in trade with third countries

The case described below was selected because it displays both of the structural elements identified in section 1.2 as the hypothesised conditions for MTIC and carousel fraud: there is a stage at which the supply is not taxed while the right to deduct input VAT is preserved, and there is artificial inflation of values within a coordinated network of formally independent participants. At the same time, it lacks the features that the literature associates with the EU legal framework, since the cross-border transactions were directed to third countries, and it lacks a missing trader, since no participant disappeared with tax collected but unremitted.

The case concerns two companies engaged in export (Exporter A and Exporter B) which, within two months, filed several tax returns claiming VAT refunds totalling some RSD 48 million, of which approximately RSD 32 million was actually paid out of the budget before the checks were completed. They purchased the goods on the domestic market from 21 suppliers and exported them, without any change in quantity or properties, to purchasers in the United Arab Emirates.

The goods the exporters purchased were not newly imported. They came from the suppliers' stocks, built up through earlier importation. Over seven years those suppliers had imported goods worth about RSD 1.1 billion and had paid VAT of some RSD 256 million on importation, which they carried in their records as a tax credit without claiming a refund while the goods remained in stock. The tax credit later used in the scheme therefore arose from an actual payment into the budget.

The tax effect arose from the combination of three steps. First, on sale to the exporters the prices of the goods were artificially inflated, by factors ranging from several tens to several hundreds relative to book value, which correspondingly inflated the VAT stated on the invoices. Second, the suppliers' liability, thus inflated, was discharged not by payment but by the tax credit accumulated earlier on importation. Third, since exports are exempt from VAT, an excess of input VAT over output VAT arose in the hands of the exporters, and that excess was converted into a refund claim. The tax stated on the domestic suppliers' invoices thus served a double function: it enabled the suppliers to use the tax credit accumulated earlier and to reduce their own liability, and it gave the exporters the basis for a refund claim against the budget, while the trade itself generated no budget revenue.

The refunds were unwarranted, therefore, not merely because nothing had been paid into the budget on those transactions. The proceedings established grounds to suspect that the trade between the suppliers and the exporters had been simulated, and that its purpose was to enable the exporters to use the VAT stated on the invoices as input VAT, which they proceeded to do by filing tax returns containing false information.

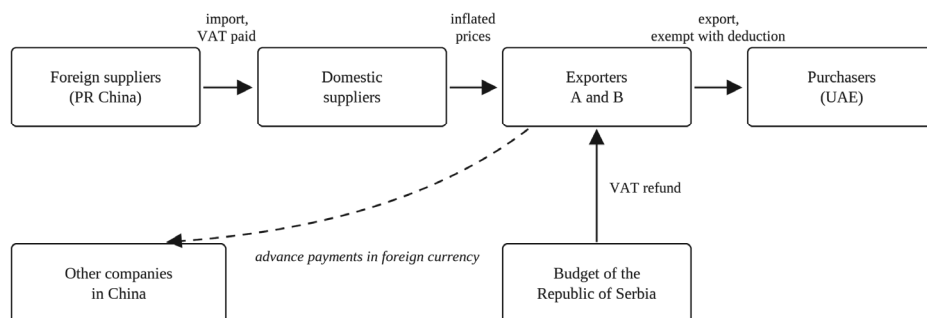
That the trade served no commercial purpose was also apparent from the exporters' trading results. Exporter A sold the goods at 3.11% below cost and Exporter B at 0.37% below cost. Trading at a loss, viewed together with the other circumstances

established, is a forensic indicator that the transactions were driven by their VAT effects rather than by trading profits.

That finding prompted further examination of possible coordination between the exporters and other formally independent participants within a scheme designed in advance. Several circumstances supported that reading, each of which taken alone may be ordinary commercial practice, but which together form a pattern. Both exporters used the same accounting agency; on all three occasions the export was carried out with the goods of both companies loaded into the same means of transport, through the same freight forwarder and the same carriers; and two purchasers were common to both exporters. During the on-site inspection the goods were not found in the leased warehouse, and the statements obtained indicated that the lessees had used the premises only briefly, so that the warehouse more probably served to support the paper trail than as a place where goods were actually kept.

Nor did the financial flows correspond to the picture of an ordinary trading transaction. Liabilities to suppliers were settled not by payment but by assignment agreements, under which the purchaser in the Emirates was to discharge obligations towards entities in China that had appeared as earlier foreign suppliers of the goods, and this without the consent of the creditors and without proper notice to the debtors. In addition, the exporters' receivables from their purchasers remained partly outstanding at the time of the inspection, which is difficult to reconcile with the conduct of a trader that genuinely intends to be paid for its goods.

The final stage of the scheme consisted in moving the extracted funds out of the country. Following the partial refund of VAT from the budget, the funds were converted into foreign currency and transferred to companies in China as advance payments for future purchases, part of those advances not being followed by any actual importation of goods. The recipients of those advances were not the same companies that had earlier appeared as foreign suppliers of the goods, so that this circumstance indicates the destination of the transferred funds but not a closed circle of the same participants. The scheme as a whole, with its flows of goods, tax and funds, is shown in Figure 1.

Figure 1: Stages of the export scheme without a missing trader

Source: authors' own work, based on data from Tax Police files.

The role played in the model by the disappearance of a trader was taken here by the accumulated import VAT credit: the suppliers set their output VAT off against that credit rather than making a new cash payment, so that no VAT was paid on that trade although the refund was claimed on it. The suspected abuse therefore concerned the transactions underlying the exporters' refund claims. Customs control confirmed that the goods had left the country, but not that the declared value was commercially justified, and it could not therefore prevent a scheme resting precisely on the inflation of that value.

3. Unwarranted VAT refunds without a cross-border element: circular trade in goods within a single jurisdiction

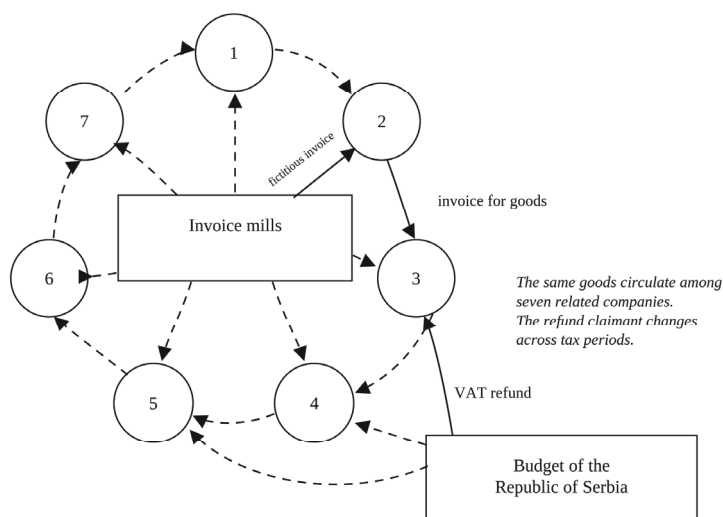
The second case study was selected to provide a contrast with the first. It exhibits participants with the characteristics of a missing trader and the actual circulation of the same goods, that is, what was absent from the first case, but it contains no transaction with a foreign party and no zero-rated supply. This makes it possible to test the proposition that unwarranted refunds can arise even where there is neither cross-border trade nor a cross-border break in the VAT chain, and to identify the elements capable of taking over that function.

The scheme operated as a closed system of seven related companies, within which the same items, bearing identical technical designations, appeared in trade repeatedly over a number of years at prices many times above actual market value. The roles within the group rotated from one tax period to the next. In one period a company would sell goods to another. Input VAT arose in the hands of the purchaser and stock was recorded, so that the purchaser filed a refund claim for that month, while the seller reduced its reported net VAT liability by claiming input VAT on a fictitious purchase invoice issued by an invoice mill. In the following period the roles changed, and the refund was claimed by another company from the same group.

The invoice mills served exclusively to document fictitious transactions. They had neither goods nor any business activity, but existed to supply fictitious purchase invoices that participants used to claim input VAT and reduce their reported net VAT liabilities. In their function within the chain they correspond to the missing trader of the model described in section 1.2, except that they also reduced their own liability by means of fictitious purchase invoices, so that no tax was paid into the budget at all. Unlike the classic model, in which the missing trader trades, collects tax and disappears, these were registered companies with no substantive trading activity: they were registered by the organisers of the scheme, documentation was drawn up for them, and their tax identification numbers were as a rule revoked (on the role of invoice mills see further Miletić, Bačić & Mišić, 2024, p. 112).

Expert examination of the goods separated the formal regularity of the documentation from the actual substance of the transactions. According to that finding, about 95% of the goods found in the warehouses of the group companies at the time of the inspection consisted of technologically obsolete and unusable waste, much of it manufactured in the Soviet era. The prices at which such goods were invoiced therefore reflected no market conditions but were set so as to inflate the taxable amount and with it the input VAT of the companies claiming refunds. The unusability of the goods is a forensic indicator that the trade served no commercial purpose and that the goods provided an apparent factual basis for the invoiced transactions.

The circulation of the goods was reconstructed through the example of a single item, an optical cable with the catalogue designation ROSENBERGER SLJ12SP-60M60R-5m, which never left the closed circle of participants and never reached final consumption. On each successive sale it served instead as the basis on which the purchaser stated input VAT, and it thereby became a means of distributing input VAT over time among the entities in the network. The movement so reconstructed is shown in Figure 2.

Figure 2: A domestic form of carousel fraud

Source: authors' own work, based on data from Tax Police files.

The solid line denotes the current tax period, in which company 2 receives a fictitious invoice from an invoice mill, sells the goods to company 3, and company 3 files a refund claim. The dashed lines denote the same relationships in other tax periods, when the roles rotate among the companies in the group.

A distinctive element of the scheme was the simulation of market demand through retail sales. Two of the companies had genuine retail outlets and cash registers, but analysis of the fiscal records showed that the goods in question were sold in minimal quantities, one or two units a month, at prices many times above their true value and most often to persons connected with the network. The retail operation thus created an appearance of genuine demand and made the prices used in the invoices exchanged within the group appear commercially plausible.

Between 2019 and 2023 the seven related companies in the scheme obtained VAT refunds totalling approximately RSD 500 million. That the scheme was sustained across five consecutive years shows that these were not isolated irregularities but a model maintained over the long term.

The role played in the classic model by the zero-rated supply was taken here by a systematic excess of input VAT, which arose because purchases at artificially inflated prices exceeded sales many times over, while the simulated retail operation made those prices appear commercially plausible. The break in the VAT chain, which the literature associates with the movement of goods between States, occurred here within a single jurisdiction, at the point at which an invoice mill stated tax that was never remitted.

The scheme also satisfies the defining characteristics of carousel fraud set out in section 1.2, in that the same goods circulate among the participants, with the functions performed in the classic model by the missing trader and by the undertaking that claims the refund at the end of the chain. The difference is that the role of claimant rotates from one tax period to the next. It is therefore a carousel in structure and in function, but not on the model that presupposes intra-Community acquisition.

Both cases show that the absence of a missing trader does not exclude VAT refund fraud. Establishing the abuse requires examination of the underlying transactions and the basis of the input VAT claims.

4. Comparative analysis: the dependence of the model's elements on the EU legal framework

The two reconstructed schemes were compared with the classic carousel model. The comparison was carried out by extracting from that model, as set out in section 1.2, and from the legal architecture described in section 1.3, the set of elements of which it is composed. Each element was then tested against both schemes and marked as present or absent, and the elements were classified according to whether they are conditioned by the EU legal framework. The result is set out in Table 2.

Table 2: Elements of the classic carousel model in the two cases analysed

Element	Classic carousel (EU)	Case 1	Case 2
A. Elements conditioned by the EU legal architecture			
Exempt intra-Community supply with self-assessment by the acquirer	yes	no	no
Movement of goods across a frontier without customs control	yes	no	no
Break in the VAT chain at the boundary of two tax jurisdictions	yes	no	no
Verification dependent on the exchange of information between two jurisdictions	yes	no	no
The question of the liability of a taxable person acting in good faith	yes	no	no

Element	Classic carousel (EU)	Case 1	Case 2
B. Elements present in all the forms analysed			
Systematic excess of input VAT over output VAT	yes	yes	yes
Artificial inflation of values within a coordinated network of participants	yes	yes	yes
No corresponding cash VAT payment in the period ⁵	yes	yes	yes
C. Elements that differ across the forms			
A stage at which the supply is untaxed with the right to deduct input VAT	yes	yes	no
Missing trader	yes	no	no
Invoice mill	no	no	yes
Circulation of the same goods through a closed circle of participants	yes	no	yes
Rotation of the refund claimant role across periods	no	no	yes

Source: authors' own work.

Group A comprises the elements conditioned by the EU legal architecture, and none of them appeared in the two cases analysed. In case 1 there was cross-border trade, but it was directed to a third country and subject to customs control of a kind that does not exist at the internal frontiers of the EU. In case 2 there was none at all. The jurisdiction of a single tax administration in both cases eliminated any exchange of information between two jurisdictions, while the question of the liability of a taxable person acting in good faith did not arise, since the participants acted in concert. The unwarranted refunds were thus obtained without any of the circumstances the literature cites as the source of the vulnerability of intra-EU trade to VAT fraud.

Group B contains three elements that appear both in the EU model and in the two cases analysed, and which together constitute the core of the model independently of the EU legal framework. The first is an excess of input VAT over output VAT, without which no refund can be claimed at all. The second is artificial inflation of values within a coordinated network of participants, which determines the amount extracted from the budget. The third is the absence of a corresponding cash VAT payment by the relevant participant in the period concerned. What the three forms

⁵ The three forms differ in what happens to the tax. The missing trader does not remit the VAT it has collected. The supplier in case 1 sets its output VAT off against an import VAT credit acquired earlier. The participants in case 2 reduce their reported net liability by claiming fictitious input VAT. All they share is that no payment reaches the budget in the period concerned; the legal nature of the operation differs in each.

share is that absence, not the way in which it comes about. None of these three conditions is tied to any legal peculiarity of the EU. All three follow from the invoice-credit mechanism itself, and exposure to this type of fraud is therefore inherent in any VAT system that applies that mechanism. This conclusion is drawn from the forms analysed and does not exclude variants in which one of the three conditions is absent.

Group C sets out the elements in which the two cases differ both from one another and from the EU model. The most significant is the first element of the group, because in case 2 there was no zero-rated supply at any stage, even though that element was hypothesised in section 1.2 as necessary. An excess of input VAT over output VAT nevertheless arose, because purchases at artificially inflated prices exceeded sales, and the simulated retail operation made those prices appear commercially plausible. What is necessary in the model is therefore not exemption as a legal category, but a state of affairs in which input VAT is left without output VAT against which to be set off. Exemption with the right to deduct produces that state most simply, since it removes the liability on the supply while preserving the right of deduction. The same effect, however, can be produced by the price structure within a network of participants. Price structure alone would not suffice, because within a closed circle one participant's sale is another's purchase. Net extraction arises only through fictitious purchase invoices, which enable the seller to reduce its reported net VAT liability without any payment, while the distribution of transactions across periods determines who claims the refund.

This also answers the question left open in section 1.2, whether the trigger for the refund at the final stage must be an export. In case 1 it was an export, but an export to a third country, which shows that this is a property of the destination principle rather than of the single market. In case 2 there was no export at all. The final stage of the chain, which Fedeli and Forte (2012) identify as economically the most significant, retains that significance even when it is not cross-border.

Customs checks in case 1 confirmed that the goods had left the customs territory, but did not establish whether the declared transaction values were commercially justified. The scheme depended on those values rather than on a failure to export the goods. The absence of a customs frontier, which the literature cites as the cause of the vulnerability of intra-EU trade to VAT fraud, therefore appears as a circumstance that facilitates execution rather than as a condition of it, and the focus of verification shifts from the movement of goods to the relationship between the declared and the actual value of the transaction.

The next two elements of Group C concern the role in which the participant appears where no tax is paid into the budget. In the EU model this is done by the missing trader; in case 1 by an ordinary supplier that trades and remains available to the authorities; and in case 2 by an invoice mill, for which disappearance is not even necessary, since it carries on no activity that could cease. The distinction rests on two criteria: the mechanism accounting for the absence of a current-period cash VAT

payment, and whether the participant carries on substantive trading activity. The role is thus functional rather than institutional, being defined by what happens to the tax in the chain rather than by the characteristics of the participant that brings it about, so that verification directed at identifying a missing trader will not detect two of the three forms described.

The last two elements of Group C determine not whether an unwarranted refund will occur but the form in which it is repeated. The circulation of the same goods through a closed circle of participants appears in case 2 as it does in the EU model, and is absent from case 1, where the goods leave the country and the chain is exhausted in a single passage. The rotation of the role of refund claimant from one tax period to the next has no counterpart either in the EU model or in case 1, and was not found in the sources reviewed, which makes it the element that distinguishes case 2.

Finally, the comparison permits a typological distinction. Instead of dividing frauds into those within and those outside the EU, the material analysed points to a typology based on the mechanism used in each scheme: non-remittance of VAT collected, the use of carried-forward import VAT credits in transactions suspected of being simulated, and claims for fictitious input VAT. The first form is developed in the literature on MTIC and carousel fraud, whereas the other two are not treated there as forms of it. The closest existing concept is the domestic missing trader fraud referred to in section 1.4, which does not, however, encompass the circulation of goods. All three nevertheless result in an unwarranted refund, although the mechanisms differ. For jurisdictions outside the EU this means that the absence of intra-Community trade does not exclude exposure. It merely alters the form in which that exposure appears.

Conclusion

In both cases analysed, unwarranted VAT refunds were obtained without any of the elements that the EU legal architecture introduces into the model. The characteristics common to all the schemes analysed are a systematic excess of input VAT over output VAT, artificial inflation of values within a coordinated network of participants, and the absence of a corresponding cash VAT payment in the period concerned. None of these follows from EU law. All follow from the invoice-credit mechanism itself, a conclusion supported by the fact that both mechanisms had been described in the VAT literature before the legal framework usually invoked to explain them had even come into being. Exemption, which the literature treats as the cause, is merely the simplest route to the first of the three conditions.

Two findings follow that supplement the model. The first is that the absence of a missing trader does not preclude VAT refund fraud. In case 1 the suppliers used carried-forward import VAT credits, whereas case 2 involved claims for fictitious input VAT supported by invoice mills. Hence the typology proposed here, based on the mechanisms used to obtain unwarranted VAT refunds rather than on EU membership. The second is that the scheme in the second case is a carousel in

structure and in function, even though it contains none of the elements that EU legislation presupposes for that concept, and that within it the role of refund claimant rotates from one tax period to the next, an element the literature reviewed does not describe. The paper's contribution is also methodological, since forensic reconstruction from tax investigation files shows which elements of the model appear in an individual scheme and which do not, something the existing non-European studies do not reveal.

The findings also have clear limits. The analysis rests on two purposively selected cases and does not support conclusions about the frequency of the phenomenon or about its share in the tax gap. The outcome of the criminal proceedings is not examined, since the paper is concerned with the structure of the scheme rather than with its classification in criminal law. The comparison, moreover, depends on the way the model has been operationalised, and a different selection of elements might yield a different picture. The factual basis derives from the findings of tax audit and pre-investigation proceedings, so that the statements made about the schemes relate to the facts as so established, while domestic legislation is cited in the version in force at the time of the cases analysed. Finally, both cases come from a single jurisdiction, so that the conclusion regarding jurisdictions outside the EU remains at the level of a demonstrated possibility rather than an established prevalence.

For tax audits, these findings imply that the verification of refund claims cannot rest on the identification of a missing trader alone, since neither of the two cases contains one. Tracing the movement of goods was decisive in case 2, whereas in case 1 it confirmed the regularity of the export without revealing the scheme, so that flows of goods, documents, tax and funds need to be combined. The three conditions identified are not proof of fraud, but they are usable as an initial test for directing verification, which should focus on the relationship between the declared and the actual value of the transaction and on the connections between formally independent participants, that is, on data not visible from the return of an individual taxable person. For tax policy, the Italian experience with mandatory electronic invoicing is instructive: cross-border fraud declined and the reduction in revenue loss was estimated at between EUR 0.6 and 1 billion in 2019 (Heinemann & Stiller, 2025, p. 195). It demonstrates the reach of measures directed at the documentary trail, but electronic invoicing alone does not establish whether declared transaction values are commercially justified, so it may need to be complemented by checks on declared values. Further research might be directed at establishing whether the rotation of the claimant role occurs in other jurisdictions as well, and whether that element can be used for system-wide detection.

References

Axel Kittel v Belgian State and Belgian State v Recolta Recycling SPRL, Joined Cases C-439/04 and C-440/04 (Court of Justice of the European Union, 6 July 2006), [2006] ECR I-6161.

Borselli, F., Fedeli, S., & Giuriato, L. (2015). Digital VAT carousel fraud: A new boundary for criminality? *Tax Notes International*, 77(8), 707-724.

Brockmeyer, A., Mascagni, G., Nair, V., Waseem, M., & Almunia, M. (2024). Does the value-added tax add value? Lessons using administrative data from a diverse set of countries. *Journal of Economic Perspectives*, 38(1), 107-132. <https://doi.org/10.1257/jep.38.1.107>

Buettner, T., & Tassi, A. (2023). VAT fraud and reverse charge: Empirical evidence from VAT return data. *International Tax and Public Finance*, 30(3), 849-878. <https://doi.org/10.1007/s10797-023-09776-y>

Bulves AD v Bulgaria, Application no. 3991/03 (European Court of Human Rights, 22 January 2009).

Bussy, A., & Tassi, A. (2025). Cross-border value added tax fraud in the European Union. *Swiss Journal of Economics and Statistics*, 161, Article 7. <https://doi.org/10.1186/s41937-025-00138-5>

Commission Implementing Regulation (EU) No 79/2012 of 31 January 2012 laying down detailed rules for implementing certain provisions of Council Regulation (EU) No 904/2010 concerning administrative cooperation and combating fraud in the field of value added tax (recast). (2012). *Official Journal of the European Union*, L 29, 13-32.

Commission Regulation (EC) No 1925/2004 of 29 October 2004 laying down detailed rules for implementing certain provisions of Council Regulation (EC) No 1798/2003 on administrative cooperation in the field of value added tax. (2004). *Official Journal of the European Union*, L 331.

Council Directive 91/680/EEC of 16 December 1991 supplementing the common system of value added tax and amending Directive 77/388/EEC with a view to the abolition of fiscal frontiers. (1991). *Official Journal of the European Communities*, L 376.

Council Directive 2006/112/EC of 28 November 2006 on the common system of value added tax. (2006). *Official Journal of the European Union*, L 347, 1-118.

Cressey, D. R. (1953). *Other people's money: A study in the social psychology of embezzlement*. Free Press.

DiGabriele, J. A., Antinozzi, H. S., Russo, C., & Kunz, P. W. (2024). Taxation and forensic accounting: Informing research and practice. *Journal of Forensic Accounting Research*, 9(1), 213-233. <https://doi.org/10.2308/JFAR-2023-021>

FATF. (2007). *Laundering the proceeds of VAT carousel fraud*. Financial Action Task Force / OECD.

Fedeli, S., & Forte, F. (2012). Border tax adjustment without borders: The EU carousel of VAT fraud. *Review of Economics & Finance*, 2(4), 55-70.

Heinemann, M., & Stiller, W. (2025). Digitalization and cross-border tax fraud: Evidence from e-invoicing in Italy. *International Tax and Public Finance*, 32(1), 195-237. <https://doi.org/10.1007/s10797-023-09820-x>

House of Lords European Union Committee. (2007). *Stopping the carousel: Missing trader fraud in the EU* (20th Report of Session 2006-07, HL Paper 101). The Stationery Office. <https://publications.parliament.uk/pa/ld200607/ldselect/ldeucom/101/101.pdf>

Jakobsson, O. (2013). *Missing trader fraud in the EU* (Master's thesis). Lund University, Faculty of Law.

Keen, M. (2007). *VAT attacks!* (IMF Working Paper WP/07/142). International Monetary Fund. <https://doi.org/10.5089/9781451867060.001>

Keen, M., & Smith, S. (2006). VAT fraud and evasion: What do we know and what can be done? *National Tax Journal*, 59(4), 861-887. <https://doi.org/10.17310/ntj.2006.4.07>

Law on Value Added Tax, *Official Gazette of the Republic of Serbia*, No. 84/2004, as subsequently amended (version in force at the time of the cases analysed).

Miletić, S., Bačić, V., & Mišić, G. (2024). Portretisanje šeme pranja novca iz poreske utaje: Da li su linije razgraničenja dovoljno jasne? [Portraying a money laundering scheme arising from tax evasion: Are the dividing lines sufficiently clear?]. In D. Malinić & S. Vučković Milutinović (Eds.), *Perspektive razvoja forenzičko-računovodstvenih kapaciteta: Izazovi za javni i korporativni sektor* (pp. 111-128). University of Belgrade, Faculty of Economics and Business, Publishing Centre.

Miletić, S., & Popović, N. (2026). The abuse of corporate vehicles in VAT refund fraud and money laundering: A forensic case study. *Ekonomika preduzeća*, 74(3-4), 141-156. <https://doi.org/10.5937/EKOPRE2604141M>

Muehlmann, B. W., Burnaby, P., & Howe, M. (2012). The use of forensic accounting experts in tax cases as identified in court opinions. *Journal of Forensic & Investigative Accounting*, 4(2), 1-34.

Naritomi, J. (2019). Consumers as tax auditors. *American Economic Review*, 109(9), 3031-3072. <https://doi.org/10.1257/aer.20160658>

Optigen Ltd, Fulcrum Electronics Ltd and Bond House Systems Ltd v Commissioners of Customs & Excise, Joined Cases C-354/03, C-355/03 and C-484/03 (Court of Justice of the European Union, 12 January 2006), [2006] ECR I-483.

Podlipnik, J. (2012). Missing trader intra-community and carousel VAT frauds - ECJ and ECtHR case law. *Croatian Yearbook of European Law and Policy*, 8, 457-472. <https://doi.org/10.3935/CYELP.08.2012.135>

Pomeranz, D. (2015). No taxation without information: Deterrence and self-enforcement in the value added tax. *American Economic Review*, 105(8), 2539-2569. <https://doi.org/10.1257/aer.20130393>

Shah, J. (2026). Using computerized information to enforce VAT: Evidence from Pakistan. *Journal of Development Economics*, 178, 103595. <https://doi.org/10.1016/j.jdevco.2025.103595>

Tait, A. A. (1988). *Value-added tax: International practice and problems*. International Monetary Fund.

Waseem, M. (2023). Overclaimed refunds, undeclared sales, and invoice mills: Nature and extent of noncompliance in a value-added tax. *Journal of Public Economics*, 218, 104783. <https://doi.org/10.1016/j.jpubeco.2022.104783>

Andrea Arbula Blecich
University of Rijeka, Faculty of
Economics and Business
ORCID: 0000-0002-1379-6410

Andrija Herceg
University of Rijeka, Faculty of
Economics and Business

FORENSIC ACCOUNTING IN DETECTING FINANCIAL MALFEASANCE AND FRAUD IN NON-PROFIT ORGANIZATIONS

UDK 657.6:343.53

DOI <https://doi.org/10.54318/czf.2026.aab.021>

Abstract

Non-profit organizations perform important social, humanitarian and environmental functions and are financed primarily through donations, membership fees and public funding. As their operations depend on the trust of donors, members and the wider public, financial malfeasance in this sector represents a serious breach of accountability and may impair the achievement of organizational objectives. Due to limited resources and complex operating environments, non-profit organizations may face challenges in maintaining effective internal controls and ensuring financial transparency. In the Republic of Croatia, where more than 50,000 associations are registered, a considerable number of organizations fail to comply with statutory financial reporting requirements. Non-profit organizations account for the smallest share of occupational fraud cases yet provide the least fraud awareness training and are most frequently fined for regulatory noncompliance. The aim of this paper is to examine the role of forensic accounting in detecting and preventing financial malfeasance in the non-profit sector. The study combines a review of the relevant literature with an analysis of documented cases of financial malfeasance in Croatian non-profit organizations and an empirical application of Benford's Law and data mining techniques to the accounting records of a selected association. The analysis identifies embezzlement, misuse of earmarked funds and fraudulent financial reporting as recurring forms of financial misconduct in Croatian non-profit organizations. Forensic accounting procedures contribute to the identification of accounting anomalies associated with potential fraud, while the empirical analysis confirms that such anomalies represent indicators rather than evidence of fraud and require further investigation. The study highlights the role of forensic accounting in supporting fraud risk assessment, strengthening internal controls and improving financial transparency and accountability in the non-profit sector.

Key words: forensic accounting; financial malfeasance; non-profit organizations; fraud detection; financial reporting

JEL Classification: M41, M42, K42

Introduction

More than 50,000 non-profit civil organizations are listed in the Republic of Croatia's Register of Associations. Forensic accounting methods and techniques can be used to verify transactions and records made by non-profit organizations whenever there is a reasonable suspicion of potential manipulation. While creative accounting is not necessarily always manipulative, it can be used for illicit purposes. The research problem stems from the weak control and oversight of non-profit organization operations in the Republic of Croatia. Given the country's population size, the number of non-profit organizations is exceptionally high. Consequently, supervising them requires significant resources and a skilled workforce. A lack of adequate oversight regarding the financial activities of non-profit organizations increases the risk of potential financial misconduct within the civil sector.

The objective of this research is to identify, based on documented cases, the principal forms of financial malfeasance affecting Croatian non-profit organizations, and to test empirically through the application of Benford's Law and complementary data mining techniques to the sales ledger of a selected association. In addition to presenting the most prevalent types of financial irregularities and their prevention, the paper's goal is to highlight the significance of forensic analysis in non-profit organizations to identify financial malfeasance and discourage potential offenders.

Unlike most existing studies, which are concentrated in the United States and United Kingdom and address non-profit fraud primarily through governance and control variables, this paper focuses on the Croatian non-profit sector, for which forensic-accounting evidence is still scarce. The paper is based on the examination and synthesis of both domestic and international literature in the domains of accounting, criminology, and forensics. This is supplemented by the examination of recorded instances of financial irregularities in Croatian non-profit organizations. Benford's Law and data mining techniques (multiple transaction analysis, transaction analysis with round numbers, checking the ratio of maximum and minimum, gap detection, and RSF test) are applied to the sales ledger of the chosen association, illustrating, rather than only discussing, how forensic accounting tools can be used.

The paper is divided into four parts. After the Introduction the paper's second chapter focuses on forensic accounting, financial malfeasance and fraud in the non-profit sector providing a literature review, occupational fraud analysis and the legal framework. In the third chapter, irregularities and empirical analysis are made using Benford's Law and data mining techniques. Finally, the fourth chapter concludes the paper with findings and direction for future research.

1. Forensic Accounting, Financial Malfeasance and Fraud in the Non-profit Sector

Non-profit organizations refer to a broad category of organizations that prioritize serving the public or communities over maximizing profits, including associations, foundations, clubs, and charity and religious organizations. Financial malfeasance and fraud in this sector are a betrayal of the trust placed in them by both donors and beneficiaries. The use of different methods for documenting and presenting items in financial statements with the intention of improving public perception is known as creative accounting. To hide losses or inflate earnings and revenue, these strategies include aggressive accounting, earnings management, income and profit smoothing, and report manipulation using fraudulent methods (Belak, 2017). Although it stays within the confines of standards, lawful creative accounting can affect stakeholders' perceptions of the organization's stability through activities like asset revaluation and expense deferral (Jović, 2014).

1.1. Literature Review

Fraud represents a significant accountability risk for non-profit organizations because fraudulent activities directly reduce resources available for achieving organizational missions and may simultaneously undermine donor confidence, public trust, and organizational legitimacy (Greenlee, Fischer, Gordon & Keating, 2007). Recent evidence confirms that research specifically addressing fraud in charities and non-profit organizations remains relatively limited and fragmented, with existing studies concentrated predominantly in the United States and United Kingdom and with substantially more attention devoted to organizational opportunities and controls than to the broader behavioral and technological dimensions of fraud (Kassem, 2026).

The empirical literature indicates that fraud risk in non-profit organizations is closely associated with the quality of organizational controls and governance. Research based on non-profit fraud cases demonstrates that internal control mechanisms are relevant both to the detection of fraudulent activity and to the magnitude of resulting losses (Holtfreter, 2008). Using a large sample of non-profit organizations for 1998–2006, supplemented by data from the redesigned 2008 IRS Form 990, Yetman and Yetman (2012) find that existing governance mechanisms improve the accuracy of reported charitable expenses. This is particularly important because fraudulent conduct in non-profit organizations is not restricted to direct theft of assets. Misclassification or misreporting of expenses can distort information used by donors, regulators, boards, and other stakeholders and may constitute an important source of fraud risk even when no direct asset diversion is initially apparent (Krishnan, Yetman & Yetman, 2006).

Governance is similarly relevant to the likelihood that fraud will be identified and formally reported. Using data from 428 non-profit organizations over the 2009–2015 period, Abu Khadra and Delen (2020) examined 21 governance-related measures

derived from IRS Form 990 filings using logistic regression and decision-tree techniques. Their results identified legal compliance, board independence, federal audit, and the use of independent accountants in compiling and reviewing financial statements as significant factors associated with the reporting of asset diversions. These findings suggest that fraud detection is influenced not only by the existence of accounting controls but also by the institutional capacity of the organization to monitor management and respond to identified irregularities. The distinction between detection and disclosure is important in this context, since research on significant asset diversions indicates that non-profit organizations do not always report detected fraud consistently or transparently (Scheetz, Wilson & Dowis 2021).

Evidence from non-profit organizations indicates that perceptions of internal control strength are associated with whistleblowing behavior, while controlling activities and monitoring important components of this relationship (Scheetz, Smalls, Wall & Wilson 2022). Qualitative evidence further suggests that whistleblowing should not be viewed as an isolated mechanism but as part of a broader anti-fraud system in which organizational controls, governance, employee attitudes, leadership awareness, and the surrounding regulatory environment interact (Arroyo Pardo, Smaili & Bensid, 2024).

Forensic accounting provides an additional layer of fraud detection by focusing on the identification, analysis, and interpretation of financial information that may indicate irregular activity. One of the clearest non-profit-specific applications of forensic data analytics is the use of Benford's Law to examine financial reports. Qu, Steinberg & Burger (2020) examined U.S. non-profit organizations using data reported on IRS Form 990 and assessed whether deviations from Benford's Law could be used to prioritize financial reports for further investigation. Although the aggregate data showed close conformity with Benford's Law, 34% of organizations did not conform at the individual level. The authors suggest that Benford-based analysis may be useful as a screening mechanism for identifying filings that warrant further examination, rather than as evidence of fraud. Analytical anomalies constitute indicators that may warrant additional procedures, but they cannot independently establish fraudulent intent or fraudulent conduct. Forensic accounting techniques are most appropriately understood as investigative tools that complement, rather than replace, internal controls, governance, audit procedures, and professional judgment.

Fraud may impose costs that extend well beyond the amount initially misappropriated. Empirical evidence from 115 non-profit organizations experiencing fraud shows that more than one quarter did not survive for at least three years following public disclosure of the fraud, with organizational size, age, and characteristics of the perpetrator associated with survival outcomes (Archambeault & Webber, 2018). Fraud detection and response should be understood not merely as mechanisms for recovering financial losses but as components of organizational accountability and resilience.

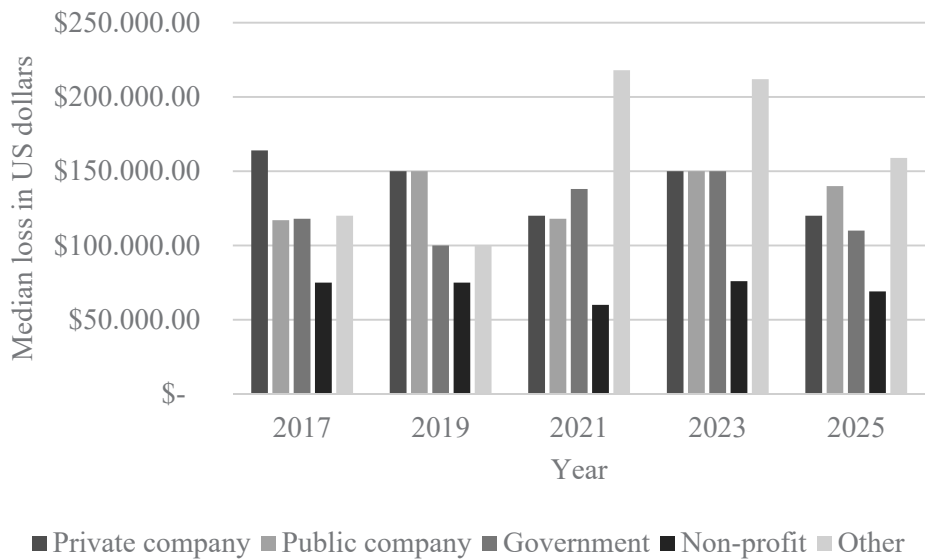
Evidence from Croatia and the wider Southeast European context remains limited compared with the international literature on non-profit fraud. Croatian research has primarily addressed non-profit accounting and financial reporting, while empirical studies directly examining fraud detection and forensic accounting in non-profit organizations remain scarce. Existing evidence nevertheless emphasizes the importance of reliable non-profit financial information for donors and organizational decision-making (Perica, Ramljak & Dragija Kostić, 2024; Rogošić, 2021), while Croatian research on forensic auditing indicates its perceived relevance for fraud prevention and detection, although not specifically within the non-profit sector (Tušek & Klikovac, 2013). Thus, the limited availability of non-profit specific empirical evidence in Croatia represents an important research gap.

The literature indicates that effective fraud detection in non-profit organizations depends on the interaction of governance, internal controls, financial reporting quality, employee reporting mechanisms, and analytical examination of financial data. Kassem (2026) confirms that existing research on charity fraud is limited and fragmented and that technological approaches to fraud detection have received relatively little empirical attention. This gap provides a basis for further research examining how forensic accounting and data-analytic techniques can complement established control and governance mechanisms in identifying potentially fraudulent transactions and financial reporting patterns within non-profit organizations.

1.2. Occupational Fraud Analysis

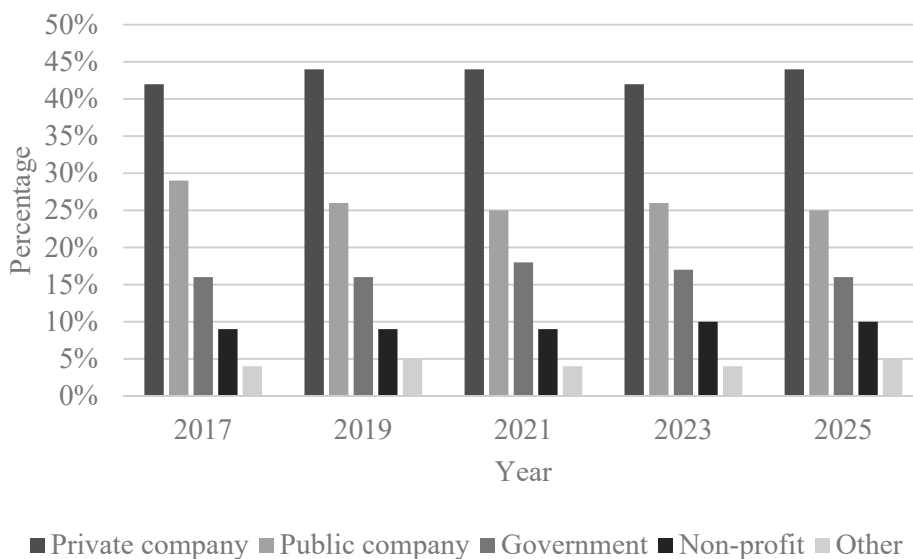
Occupational fraud is formally defined as the use of one's occupation for personal enrichment through the deliberate misuse or misapplication of the employing organization's resources or assets (ACFE, 2026). Non-profit organizations recorded the lowest median losses associated with occupational fraud among all sectors during the last five-year period (shown on Graph 1), according to a report by the Association of Certified Fraud Examiners based on cases from 138 countries and 22 industry categories.

Graph 1: Median loss of organizations that are victimized by occupational fraud (2017 – 2025)



Source: Author`s work based on Occupational Fraud 2026: A Report to the Nations, ACFE

Private companies were the most frequent victims of business fraud during the observed period, accounting for 44% of cases (shown on Graph 2) and suffering an average loss of \$120,000. Public companies were the next most frequent victims, representing 25% of cases with losses amounting to \$140,000, followed by the government and its institutions, which accounted for 16% of total cases and losses of \$110,000. Non-profit organizations represented the smallest share of cases amounting to 10% and experienced losses nearly half the size of those seen in other types of organizations, with an average loss of \$69,000 (ACFE, 2026).

Graph 2: Percentage of cases in which organizations were targets of occupational fraud (2017 - 2025)

Source: Author's work based on Occupational Fraud 2026: A Report to the Nations, ACFE

2. Irregularities and Empirical Analysis in Croatian Non-profit Organizations

Non-profit organizations face significant difficulties due to financial fraud, which substantially jeopardizes their ability to fulfill their missions and maintain stakeholder trust. Three typical types of financial irregularities are identified by documented incidents from Croatian non-profit organizations: embezzlement of funds, misuse of funds and fraudulent reporting. Funds available to and collected by non-profit organizations can be subject to financial malfeasance, ranging from methods of siphoning money out of the organization to simple, unsophisticated cash theft committed by individuals with direct access to cash inflows or available funds (Huzanić, 2015). Unlike the embezzlement of non-profit organization funds which typically involve siphoning money out of the organization or stealing cash, misuse entails using available funds for other ends. Through fraudulent reporting, non-profit organizations create a misleading impression of their operations, whether for internal stakeholders or external users of financial statements. Internal fraudulent reporting pertains to all accounting records and other documentation used within the organization for the needs of management and the internal organizational structure. In the context of non-profit organizations, external fraudulent reporting refers to all financial statements submitted to competent authorities such as balance sheets and

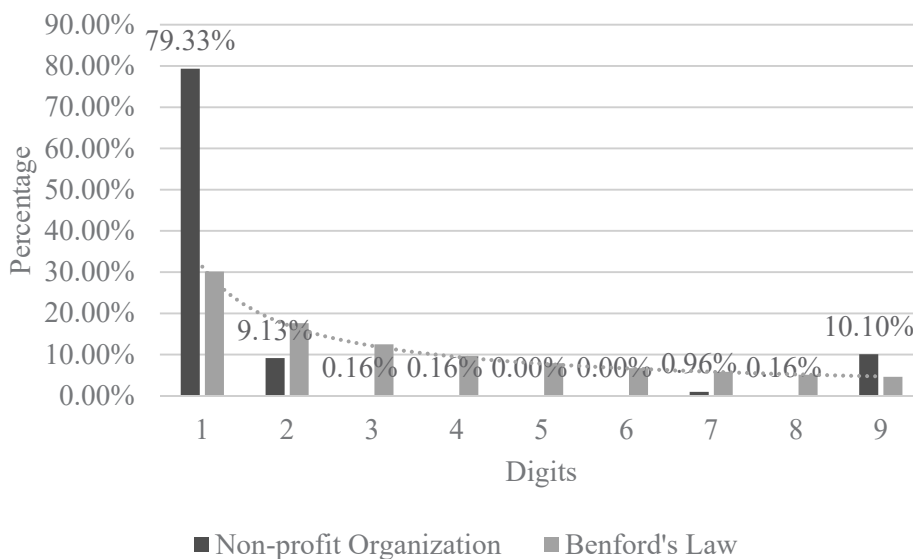
statements of income and expenditure filed on an annual or semi-annual basis (Vašíček, 2009).

2.1. Benford's Law and Data Mining Techniques

Benford's Law concerns the probability of the occurrence of specific digits in unbiased numerical datasets and is considered a tool for detecting business fraud. Through this distribution, Benford demonstrated that certain digits appear more frequently than others in datasets drawn from sources that adhere to non-mathematical guidelines (Kopal, Nemeth & Leinweber, 2019). Mark Nigrini made the most significant strides in applying Benford's Law to the detection of accounting fraud. He emphasized that when a person manipulates numbers, the choice of numbers is not random and is unlikely to follow Benford's distribution. In addition to Benford's Law, Beneish's M-score model and Data mining techniques are applied in forensic accounting. Beneish's M-score model, which is based on American accounting standards (GAAP) and the impossibility of detecting manipulations that persist over multiple consecutive periods, serves as an indicator of financial report manipulation (Beneish, 1997). Data mining is the search of massive data sets to find previously undetected abnormalities, such as double posting of accounts, transactions with round figures, or deviations in the ratio of the highest sums, is a complement to forensic approaches (Nigrini, 2011).

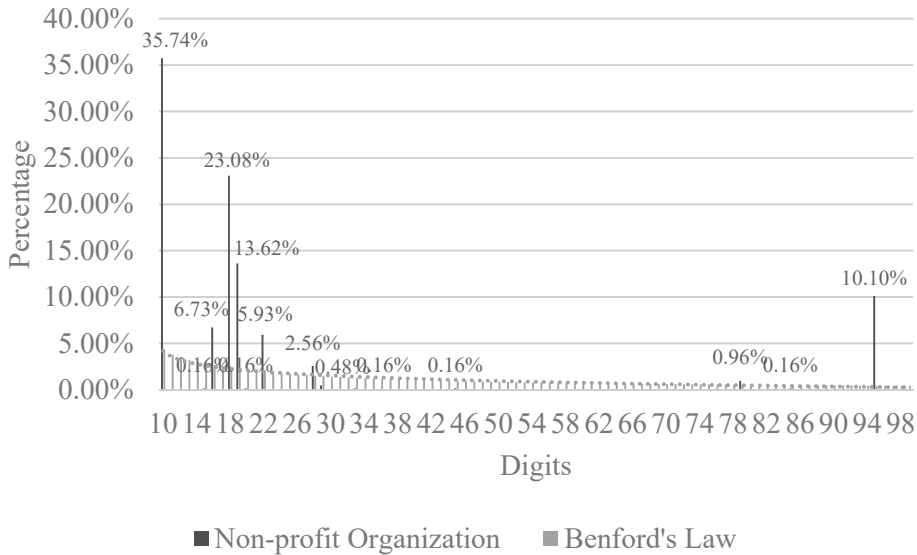
2.2. Application of Forensic Accounting in the Non-profit Sector

For the purposes of this paper, an analysis will be conducted using the example of a non-profit organization to determine whether there are indications of fraud. The non-profit organization analyzed in this paper is registered as an association. It is primarily funded through membership fees, membership contributions, and donations, for which it issues invoices and records them in the sales ledger. An analysis using Benford's Law for the first digit (shown on Graph 4) and first two digits (shown on Graph 5) was conducted based on the sales ledger, which contains 624 recorded invoices from the calendar year. The results obtained deviate significantly from the standard Benford distribution.

Graph 3: Visualization of the application of Benford's Law with the first digit to a non-profit organization

Source: Author's work based on the non-profit organization's sales ledger

It is evident that the non-profit organization's sales ledger contains no invoices starting with digits 5, or 6. The association issued the highest percentage of invoices starting with digit 1 (79.33%) out of which the most starting with digits 10 (35,76%), followed by invoices starting with the digit 9 (10.10%), all of which start with digits 95. Regarding the remaining leading digits, the association issued a total of 9.13% of invoices starting with the digit 2, 0.96% starting with 7, and 0.16% starting with the digits 3, 4 and 8. Further analysis of the sales ledger reveals that the association predominantly received membership fees, membership contributions, and donations; these items are precisely the reason for the uneven distribution and the deviation from the standard Benford distribution. The association issued the highest number of invoices for membership contributions starting with digits 1, 2 and 9, and fewer starting with digits 3, 4, 7 and 8.

Graph 4: Visualization of the application of Benford's Law with the first two digits to a non-profit organization

Source: Author's work based on the non-profit organization's sales ledger

The chi-square goodness-of-fit analysis indicates a substantial deviation from Benford's Law for both the first digit and the first two digits. For the first digit, $\chi^2 = 850.03$ ($df = 8$), compared with the critical value of 15.507, while for the first two digits, $\chi^2 = 4,836.43$ ($df = 89$), compared with the critical value of 112.022, at the 5% significance level. A chi-square goodness-of-fit test was applied to formally assess this conformity, testing the null hypothesis (H_0) that the observed leading-digit distribution follows Benford's Law against the alternative hypothesis (H_1) that it does not. Since both calculated values substantially exceed their respective critical values, H_0 is rejected for both the first-digit and first-two-digit distributions, providing strong statistical evidence that the observed distribution differs significantly from the expected Benford distribution. However, this statistical deviation alone should not be interpreted as evidence of fraud or data manipulation, it indicates only that the dataset departs from the pattern expected under Benford's Law and warrants further examination.

The analysis is based on 624 observations, which are below the often-recommended sample size of 1,000 for Benford's Law analysis. The results should be interpreted with some caution, as smaller samples may be more sensitive to deviations caused by the characteristics of the dataset. Nevertheless, the magnitude of the observed chi-square statistics indicates a very substantial departure from the expected Benford distribution.

A review of the sales ledger revealed that the association issued invoices with a maximum amount of €338.00, a second highest amount of €295.00, and a minimum amount of €45.00. Calculating the ratio between the highest and lowest invoice amounts yields a figure of 7.51, indicating a discrepancy that warrants further investigation. Further analysis established that the lowest amount represented a small membership fee, while the highest was a combined donation from several members; this explains the magnitude of the ratio between the funds paid into the organization's account. Calculating the RSF test, the ratio between the highest and second-highest amount yields a result of 1.15. A further review of the outgoing invoices register confirmed that both invoices were issued to members that had made donations and membership contributions to the association. No gap detection was observed, and no duplicate transactions exist. Many invoices were issued for round amounts; further examination revealed that individual project participants were paying their own participation fees, resulting in many invoices with round figures.

The fundamental principle of forensic analysis, which states that statistical indicators serve as the foundation for additional verification rather than independent evidence of irregularities, was confirmed by an additional check, which found that the aforementioned deviations originate from a legitimate difference between the amount of member contributions and the individual donation of a legal entity (Nigrini & Wells, 2012). Given all the above, it can be concluded that the results obtained are consistent and in line with Qu et al. (2020) examination of non-profit organizations.

Conclusion

The term "creative accounting" refers to a variety of methods, ranging from the intentional falsification of financial data, which includes embezzlement and the falsification of income or expenses, to the lawful modification of accounting practices within the parameters of standards. Using quantitative methods such as Benford's Law and data mining, together with an understanding of the motivation, opportunity, justification, and capabilities of the perpetrator, forensic accounting can help identify warning signs and anomalies associated with such irregularities and can support further investigation - although, on its own, it cannot establish that fraud has occurred; that determination requires additional evidence and, ultimately, a formal investigative or judicial process.

Embezzlement is the most widespread form of fraud, involving trusted individuals who divert funds through sophisticated schemes such as forged documents or inflated expenses. Misappropriation of funds occurs when resources intended for charitable purposes are diverted to cover personal expenses. Fraudulent reporting undermines accountability and transparency - core values of any non-profit organization. It involves manipulating financial data or exaggerating achievements to attract donors or meet grant requirements, which can mislead stakeholders and damage the organization's credibility. A single non-profit organization and a

comparatively small sample of 624 invoices from a single calendar year were used in the empirical application of Benford's Law and data mining techniques; this is below the suggested minimum sample size for a more dependable application of the Benford distribution and restricts the ability to generalize the results to other organizations or time periods.

The structural characteristics of the association's revenue account for the observed deviations from the Benford distribution in this instance rather than manipulation, which further demonstrates that forensic quantitative methods only serve as a guide for additional research rather than proving fraud on their own. To more accurately establish a cause-and-effect relationship between detected abnormalities and actual financial irregularities, future research should include a larger dataset covering a longer period from many non-profit organizations and integrate diverse forensic approaches. Benford's Law and the complementary forensic accounting techniques applied in this paper are best understood as tools for identifying anomalies and warning signs that justify further scrutiny, not as methods capable of independently proving that fraud has occurred; any such conclusion requires corroborating evidence obtained through additional investigative procedures.

References

- Abu Khadra, H., & Delen, D. (2020). Nonprofit organization fraud reporting: Does governance matter? *International Journal of Accounting & Information Management*, 28(3), 409–428. <https://doi.org/10.1108/IJAIM-10-2019-0117>
- Archambeault, D. S., & Webber, S. (2018). Fraud survival in nonprofit organizations: Empirical evidence. *Nonprofit Management and Leadership*, 29(1), 29–46. <https://doi.org/10.1002/nml.21313>
- Arroyo Pardo, P., Smaili, N., & Bensid, S. (2024). Removing barriers to whistleblowing at nonprofit organizations through employee empowerment. *Accounting Perspectives*, 23, 267–300. <https://doi.org/10.1111/1911-3838.12332>
- Association of Certified Fraud Examiners. (2026). *Occupational Fraud 2026: A Report to the Nations*, retrieved on August 15, 2026, from: <https://www.acfe.com/fraud-resources/report-to-the-nations>
- Belak, V. (2017). *Falsification of financial statements, fraud and forensic accounting*. Belak Excellens d.o.o.
- Beneish, M. D. (1997). Detecting GAAP violation: Implications for assessing earnings management among firms with extreme financial performance, *Journal of accounting and public policy*, 16(3), 271-309.
- Greenlee, J., Fischer, M., Gordon, T. P., & Keating, E. K. (2007). An investigation of fraud in nonprofit organizations: Occurrences and deterrents. *Nonprofit and Voluntary Sector Quarterly*, 36(4), 676–694. <https://doi.org/10.1177/0899764007300407>

Holtfreter, K. (2008). Determinants of fraud losses in nonprofit organizations. *Nonprofit Management and Leadership*, 19(1), 45–63. <https://doi.org/10.1002/nml.204>

Huzanić, B. (2015) Kreativno računovodstvo i utjecaj na poreze. *Računovodstvo, revizija i financije* (RRIF), br. 5/15. 26-33.

Jović, D. (2014). Loši krediti i kozmetičko računovodstvo, *Tranzicija*, 16(33), 30-46. <https://hrcak.srce.hr/127837>

Kassem, R. (2026). Charity fraud in focus: A systematic analysis of typologies, drivers and countermeasures. *Journal of Accounting Literature*. <https://doi.org/10.1108/JAL-05-2025-0268>

Kopal R., Nemeth T., & Leinweber G. (2019). Primjena Benfordovog zakona u otkrivanju računovodstvenih prijevара, *FIP - Financije i pravo*, 7(2), 7-22. <https://hrcak.srce.hr/238085>

Krishnan, R., Yetman, M. H., & Yetman, R. J. (2006). Expense misreporting in nonprofit organizations. *The Accounting Review*, 81(2), 399–420. <https://doi.org/10.2308/accr.2006.81.2.399>

Nigrini, M. & Wells, T. (2012). Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection <https://doi.org/10.1002/9781119203094>.

Nigrini, M. (2011). Forensic Analytics: Methods and Techniques for Forensic Accounting Investigations. <https://doi.org/10.1002/9781118386798.ch1>

Perica, I., Ramljak, B., & Dragija Kostić, M. (2024). The effects of information from NGOs' financial statements on charitable contributions: Evidence from Croatia. *Ekonomski pregled*, 75(2), 125–146. <https://doi.org/10.32910/ep.75.2.1>

Qu, H., Steinberg, R., & Burger, R. (2020). Abiding by the law? Using Benford's Law to examine the accuracy of nonprofit financial reports. *Nonprofit and Voluntary Sector Quarterly*, 49(3), 548–570. <https://doi.org/10.1177/0899764019881510>

Rogošić, A. (2021). Conversance with accounting in non-profit organisations and its impact on decision-making and financial planning. *Zbornik radova Ekonomskog fakulteta Sveučilišta u Mostaru*, 27, 9–24. <https://doi.org/10.46458/27121097.2021.27.9>

Scheetz, A. M., Smalls, T. D. W., Wall, J., & Wilson, A. B. (2022). Perception of internal controls helps explain whistleblowing. *Nonprofit and Voluntary Sector Quarterly*, 51(4), 759–782. <https://doi.org/10.1177/08997640211017665>

Scheetz, A. M., Wilson, A. B., & Dowis, W. B. (2021). An investigation of nonprofit reporting of significant diversions of assets. *Journal of Forensic Accounting Research*, 6(1), 176–206. <https://doi.org/10.2308/JFAR-2020-018>

Tušek, B., & Klikovac, A. (2013). Analysis of possible forensic audit models in the Republic of Croatia. *Ekonomski pregled*, 64(2), 101–122. <https://hrcak.srce.hr/100647>

Vašiček, D. (2009). Računovodstvo i financijsko izvještavanje u funkciji upravljanja u javnom sektoru, *Hrvatska i komparativna javna uprava*, 9.(2.), 393-420. <https://hrcak.srce.hr/135758>

Yetman, M. H., & Yetman, R. J. (2012). The effects of governance on the accuracy of charitable expenses reported by nonprofit organizations. *Contemporary Accounting Research*, 29(3), 738–767. <https://doi.org/10.1111/j.1911-3846.2011.01121.x>

Goran Petković
University of Belgrade, Faculty of
Economics and Business
ORCID: 0000-0001-7304-9751

RISK ASSESSMENT OF SERVICE PROVIDERS' BUSINESS OPERATIONS - A COMPARATIVE OVERVIEW OF DIFFERENCES IN APPROACH USING THE EXAMPLE OF CEFTA MEMBERS

UDK 338.48:366.5

DOI <https://doi.org/10.54318/czf.2026.gp.022>

Abstract

The freedoms of movement of people, goods, services and capital are key achievements of European integration, the basis for the prosperity of the community as a political and economic organization. In such an economic space, key market policies are focused on consumer protection and competition protection. Both are based on transparent reports and procedures that should provide insight into business operations and compliance with regulations in that business. The protection of consumers of services is a priority for regulators, also in the case of tourism authorities in CEFTA members. However, consumer protection mechanisms differ from each other in each market, and they also differ from European standards that are based on monitoring the volume of business activities in the business books of tourism service providers. The purpose of this research is to contribute to the answer on the question why regulators of tour operators and travel agencies in CEFTA members do not base their consumer protection clauses on accounting data of business entities in their markets?

Key words: *business activity, business risk, consumer protection, tourism, CEFTA, harmonization*

JEL Classification: *L83, M38, F53*

Introduction - EU Traveller Protection

The basic business rules prescribed by the general legislative framework proved to be insufficiently specific and effective in the case of specialised products and services, including package travel. This led to the adoption in 1990 of a separate EU Directive on packaged travel, packaged holidays and packaged tours (Council Directive 90/314/EEC, 1990). Approximately a decade and a half later, this framework was significantly improved in response to the growth of online sales of tourist services. In 2015, the revised framework extended protection to both purchasers of packaged travel and purchasers of linked travel arrangements (DIRECTIVE (EU) 2015/2302). Its aim was to clarify relevant terms and definitions and to strengthen consumer protection by preventing the avoidance of obligations through broad interpretations of the general legal framework. One important

clarification concerned the obligations to provide reimbursement and repatriation of travellers who had paid for their trip in advance and were affected by the insolvency of a tour operator before or during the trip. The Directive established an obligation to provide financial protection for travel and referred to guarantee funds or insurance company, while leaving the specific arrangements for such guarantees to the Member States and the market. These rules were further refined in 2018 when airlines were classified as tour operators where accommodation and other tourist services were sold through their websites or branches (Fact sheet, 2018). The amendments also regulated the protection of passengers in the event of airline insolvency, applying principles similar to those previously established for users of tour operator services.

The COVID19 pandemic has caused numerous travel disruptions and cancellations, creating significant financial pressure on both tour operators and institutions providing financial guarantees against insolvency and causing high interest in academic literature, so that 29,22% of researches on financial protection of travellers were published between 2019-2022 (Almansoori et al., 2024). Consequently, emerged market situation highlighted the need of regulators to review the effectiveness of the existing Directive (European Commission, 2021). The new Directive adopted in 2026 removed the potentially confusing concept of linked travel arrangements and, in order to strengthen travellers protection in the event of insolvency, shortened the refund period for cancelled trips to 6 months (with the exception provided in Art.17(6)), enabled member states to establish limits on advance payments for package travel, introduced the possibility of additional insurance through a protection fund financed by tour operators, and introduced more detailed rules on vouchers issued by tour operators during periods of mass cancellations of trips and large-scale refund obligations (DIRECTIVE (EU) 2026/1024).

Even before adoption of the new Directive, a number of countries had taken steps to strengthen consumer protection. In Germany, for example, Der Deutsche Reisesicherungsfonds has been operating since 2021. It is financed by private members and provides compensation in the event of tour operator insolvency. (<https://drsf.reise/>).

1. Insolvency Protection Mechanisms for Tour Operators in European Economies

The new 2026 Package Travel Directive has not yet been transposed into the legislation of European countries, although many countries already apply some of the solutions introduced by the new Directive as general rules. Based on a review of existing mechanisms for protecting travellers against tour operator insolvency and, for comparative purposes, the framework established by the 2015 Directive, the legislation of European countries can be categorized into four groups.

Travel guarantee based on central guarantee fund with Germany as the main example. The mandatory Central Guarantee Fund (Deutscher Reisesicherungsfonds) operates under the state supervision and has a target capital of EUR 750 ml (RSG, 2021). The legal basis for this system is the German Civil Code, i.e. Bürgerliches Gesetzbuch (BGB), §§ 651r–651t (BGB, 2002). The deadline for reaching the target capital is 31.10.2027, by which time the state treasury guarantees the payment of damages and takes over claims for compensation. The fund is filled with payments from tour operators so that the collateral reaches 5% of their sales and cannot be less than 1% of their annual turnover. After the target capital or the deadline has been reached, the state will withdraw from financing the Fund's obligations. Although academic analyses of the new measure's impact on tourism in Germany are not yet available, one must consider the legal community's fatigue regarding the constant alignment of national legislation with the European legal framework, as well as warnings about the resulting issue of legal uncertainty (Gago, Hruschka, 2022).

Travel guarantee based on private guarantee fund, with the Netherlands as notable example. Following the transposition of the 2015 Directive into the Dutch Civil Code, the new Article 513a, par. 1, Book 7, stipulates: „1. The organiser shall take whatever measures are needed to ensure that if he is unable or no longer able to fulfil his obligations towards the traveller owing to financial insolvency, his obligations will either be assumed by someone else, or the price of the package - or part thereof, if some of the travel services have already been used - will be reimbursed. If the package travel contract includes the carriage of passengers, the measures shall provide for the travellers' repatriation (NL, 2018). Without prescribing a specific mechanism for fulfilling this obligation, the state left it to business operators to determine the appropriate form of financial protection. To manage this risk, tour operators have established private guarantee funds such as STO Garant (217 members according to www.sto-garant.nl, 2026), VZR Garant (460 members) and CGTO which focuses on small and niche market tour operators. The funds themselves establish participation requirements and contributions. For example, some require 5 EUR per passenger (boatbiketours.com, 2026). The funds commonly hold collected contributions in escrow accounts, although other arrangements are also possible. Thus, current Dutch regulations regarding the protection of tour operator clients do not alter the stance on whether financial security mechanisms should be organized by the government or the travel sector, continuing to place full confidence in private initiative (Faure, Weber, 2013).

Travel guarantee based on insurance or bank guarantee, model applied in many countries in the region, like Austria, Greece, Croatia, Romania, but also some other like France. Under this approach, tour operators are required to obtain financial guarantee from bank or insurance company, usually as the part of licencing process. France requires tour operators to register with a financial guarantee against insolvency, issued by credit or insurance institution established in the EU (FR, 2017-1717). The amount of the guarantee is not prescribed, but there is an obligation to refund unused payments and provide repatriation where the trip has already

commenced. Similarly, Romania requires applicants to submit one of several instruments that can be used for licensing purposes: „letters of bank guarantee, insurance policies, a package travel guarantee fund or other legally established guarantee instruments“ (Guvernul, 2018). In addition, a minimum capital requirement of RON 25,000 (approximately EUR 4,760) applies to travel organizers, while the amount of the financial guarantee itself is not prescribed. Similarly, Greece and Croatia, both of which are geographically and institutionally close to CEE region, require insurance or bank guarantee covering repatriation and refund of advance payments, but do not prescribe a statutory minimum amount (Presidential Decree, 2018; Zakon, 2017). Austria, by contrast, prescribes a minimum amount of insurance policy or bank guarantee or public law corporate guarantee. The required amount is highest 13.000Eur or 18% of the annual turnover or 50% of the turnover in the peak month (PRV, 2018). Such regulation of the minimum insured amount is less typical of market-regulated economies and therefore reflects a more administrative (bureaucratic) approach to the regulation of market relations. This approach may also have influenced countries in the region, whose legislative frameworks have historically been influenced by Austrian and more broadly Austro-Hungarian legal traditions.

Travel guarantee based on parallel insurance of travel agencies and air carriers (hybrid model), as applied in Great Britain. With regard to the financial protection of tour operators, the applicable 2018 regulation recognizes several instruments for protection against insolvency. Where a bond is used as the guarantee instrument, the amount of bond must be calculated within a range from 25% to a maximum of 100% of the planned travel package turnover during the business year (The Package Travel, 2018). No specific amount is prescribed for an insurance policy, on the assumption that the insurer will assess the risk and determine the amount necessary to cover all costs of refunding advance payments. For travellers purchasing tourist services through an airline, a specific form of traveller protection applies through ATOL (Air Travel Organiser's Licence) (The Civil Aviation, 2012). ATOL was established in 1973 and subsequently upgraded in 2012 and 2018. It represents a government-regulated financial protection scheme designed to protect travellers against airline insolvency, whether the traveller is already on holiday or has not yet commenced the trip (Conway, 2019). However, lawyers in Great Britain showed some doubts towards this “original solution” for consumer protection (Deacon, 2025).

Based on the above review and comparison of regulations governing traveller protection against tour operator insolvency, the most common model of risk protection consists of financial guarantees provided by banking and insurance institutions. The presented analysis of European practice regarding the protection of travellers using tour operator services establishes a framework for further analysis of the same phenomenon in CEFTA countries, which will be the subject of research in this study.

2. Research Methodology

The subject of the research will be an analysis of the legal framework in CEFTA member states governing the protection of customers of travel organizers. Research into the protection of package travel consumers is not unknown in the literature (Grant, 1992; Chen, 2021), but this phenomenon has not yet been investigated within the context of the CEFTA market. An exception from this analysis will be Moldova, which has abolished the licensing requirements for travel agencies and tour operators in line with the European approach of retaining licensing requirements only where there is a justified consumer protection interest. Consequently, the abolition of the licensing requirement also eliminates several criteria previously associated with obtaining a license, including the obligation to provide a travel guarantee. Moldova is therefore excluded from the subsequent analysis of CEFTA members.

The aim of this research is to identify any differences that may exist between consumer protection for users of tour operator services in CEFTA countries and that of the European Common Market. In this context, two research questions can be defined: RQ 1. What are the key differences regarding passenger protection in CEFTA members compared to European members? RQ2. What are the causes of such differences?

The research was conducted in two phases, first comparative analysis of the legal regulation and after that, in second round, discussion with the regulators, which is well known research design (Eisma, Christie & Hershman, 2005). The first phase was a legal framework review. The aim of this research is to establish the facts regarding how tour operator service users are protected, specifically, which instruments are used, against which risks as well as how the level of risk is assessed in various CEFTA members. In the second phase, interviews and discussions were conducted with representatives of the institutions in charge for the regulation of tour operators' activities in each CEFTA member. The aim of the second phase of the research was to determine the rationale behind the observed regulation of tour operator activities and to assess potential weaknesses as a basis for the future improvement of the existing legislative framework, as previously described in the other situation involving the harmonization of the tourism regulator's operations (Charles, 2019).

3. Travel Guarantees in The Event of Tour Operator Insolvency in CEFTA Members

CEFTA members have generally adopted the most frequent model of mandatory financial guarantees, and in implementing this model, have largely followed the Austrian approach. This means that the financial guarantee mechanism is supplemented by statutory minimum guarantee amount. However, CEFTA members have not adopted identical solutions in this regard.

The following Table 1 shows comparative of mechanisms used to protect travellers against the insolvency of tour operators in CEFTA members. Several limitations of the analysis should be noted. First, solutions contained in draft legislation in some member parties are not included. Second, in the case of Bosnia and Herzegovina, the analysis does not cover solutions applicable at the cantonal level or in the Brčko District, instead, it only considers the legislation at the level of the two entities of BiH.

Table 1: Overview of instruments for protecting travellers against the insolvency of tour operators in CEFTA

	Law	Bylaw act	Instrument(s) of insolvency protection	Minimal value	Purpose
Montenegro	Article 17 The law on Tourism and hospitality OG 2/18, 04/18 and 13/18	/	Irrevocable bank guarantee Insurance policy Monetary deposit	15.000 Eur (tour operator) 7.000 Eur (agent) In future: % of total revenue from P/L statement for previous year	Compensation of damage (failure to fulfil contract, accommodation, food, return to place of departure)
Serbia	Art.58 Law on Tourism, Official Gazette of the Republic of Serbia 17/2019	Rules on the type, amount and conditions of the travel guarantee, the amount of the deposit depending on the license category, the method of activation, the purpose of the funds, as well as other conditions that must be met by the travel organizer, Official Gazette of Republic of Serbia, 82/2023	Bank guarantee Insurance policy Monetary deposit	Category/Passengers A10: 50.000€ A15: 100.000€ A20: 250.000€ A25: 300.000€ A35: 350.000€ A40: 400.000€	Accommodation, food and repatriation, refund of advance payment
Albania	Art.64 Law No. 93/2015 Law No.30/2024	DCM 710/2016	Insurance contract	Liabilities resulting from failure to fulfil obligations	Liabilities resulting from failure to fulfil obligations
North Macedonia	Art.31-33 Law on Tourist Activity	/	Insurance contract	A: 100.000€ B: /	Insurance of persons and, refund for undelivered services, repatriation
Kosovo	Art.11 Law No. 08/L 074 on Tourism Activity	Administrative Instruction (MIET) No. 04/2023 On Tourism Registry	Irrevocable Bank guarantee Insurance policy	Not prescribed	Damage, accommodation, food and return to the place of departure

Table 1: Overview of instruments for protecting travellers against the insolvency of tour operators in CEFTA

	Law	Bylaw act	Instrument(s) of insolvency protection	Minimal value	Purpose
Republic of Srpska	Art.34, Art.35 Law on Tourism, Official Gazette 45/2017 and 16/2023	/	Irrevocable, unconditional and payable on demand bank guarantees Insurance policy	100.000 KM (50.000€) – organizing travels in and out of country 30.000 KM (15.000€) organizing travels in country A special bank guarantee in the minimum value of half of the total cost of the excursion	Refund of advanced payment; accommodation, food and return to the place of departure; damage
Federation of Bosnia and Herzegovina	Art.23 Law on Tourist Activity, Official Gazette of the Federation of Bosnia and Herzegovina, 2009/32–	/	Bank guarantee Insurance policy Cash deposit	A guarantee for each travel arrangement	Refund; Return to the place of departure

Source: Author`s research

Several conclusions can be drawn from this analysis.

Variations in risk coverage across CEFTA members. The scope of risks for travel agency customers is not uniform across CEFTA members. Nevertheless, all systems cover the risk of tour operator insolvency, particularly the obligation to refund payments where the contracted service have not been provided and to repatriate travellers to the place of departure when the journey has already commenced (Federation of Bosnia and Herzegovina). Most of regulations examined also cover the costs of food and necessary accommodation incurred by travellers, until their return to the point of departure (Montenegro, Serbia, Kosovo, Republic Srpska). In Kosovo, additional compensation for damage is provided, and in North Macedonia there is additional personal insurance, while in Albania, coverage of all costs incurred due to non-fulfilment of obligations is provided.

Instruments for risk coverage vary across CEFTA members. The most commonly used instruments of guarantee are bank guarantees and insurance policies issued by local financial institutions, although significant variations exist. Albania and North Macedonia rely exclusively on insurance and do not provide for bank guarantees. Montenegro, Serbia and the Federation of Bosnia and Herzegovina, also allow cash deposits, as an alternative form of financial security. Such deposits are not generally envisaged under European rules for the protection of travel services users. This raises an important question concerning their effectiveness: a deposit held by the tour operator is not directly accessible to the traveller in distress as required by Directive (EU) 2015/2302, Article 17(4). Furthermore, in the event of bankruptcy, deposit accounts are frozen and become part of the pool of assets available to creditors on a first-come, first-served basis, where travel service users would likely be left behind, unprotected. More sophisticated forms of escrow or pledge accounts for cash deposits have not yet been widely considered. Nevertheless, the availability of cash deposits has proved necessary in some CEFTA markets because insurers are often reluctant to issue policies, arguing that the risks associated with tour operator insolvency are difficult to assess. In Serbia, this has reportedly been an issue for some operators, while in Kosovo insurers refuse to provide coverage altogether, contributing to the suspension of the licensing process. In Montenegro and Kosovo, the legislation explicitly requires bank guarantee to be irrevocable, while legislation of Republic Srpska, additionally requires guarantees to be *unconditional and payable on demand*.

Adoption of fixed minimum risk coverage. A substantial number of CEFTA jurisdictions prescribe a minimum amount of travel financial security as a condition for conducting tour operator activity. Except in the Federation of Bosnia and Herzegovina, this requirement is generally incorporated into the licensing procedure. In the Federation of Bosnia and Herzegovina, licensing has been delegated to the cantonal authorities and in most cases, the financial guarantee is likewise a mandatory licensing requirement. Kosovo does not prescribe a fixed minimum amount of a financial security, while Albania requires coverage “liabilities resulting from failure to fulfil obligations”, thereby leaving the assessment of the relevant

business risk largely to the insurer. Similarly, the legislation of Federation of Bosnia and Herzegovina, provides for a “a guarantee for each travel arrangement”, suggesting a comparable approach.

Open issues of fixed risk coverage. The question arises as to why fixed minimum amounts of financial security are prescribed in so many cases. This approach has several evident shortcomings. First, travel agencies differ substantially in terms of turnover and business volume. Consequently, a fixed guarantee amount cannot provide an equivalent level of protection for customers of all agencies. Indeed, a uniform minimum amount may provide the least protection for customers of agencies with the highest turnover. Serbia has attempted to address this problem by differentiating the required guarantee according to the size of the agency and the number of passengers. However, agency turnover varies from year to year and, in particular, from season to season. This makes the gradation procedure of the amounts unsustainable, particularly in cases of a sudden increase of turnover beyond the level covered by the existing travel guarantee. A sudden increase in turnover is particularly relevant because it may correspond to a simultaneous increase in business risk. Conversely, setting fixed guarantee amount at a high level, with the objective of providing greater protection to travellers, may disproportionately affect small agencies by imposing conditions that are difficult or even impossible for them to meet when applying for a licence. In response to these financial barriers, small tour operators may establish business alliances in which several small tour operators obtain a single travel guarantee through one legal entity, share the associated costs and market travel packages as a product of that legal entity while acting as its agents. Although such arrangements may provide practical solution to the financial burden imposed by mandatory guarantees, they can also reduce the transparency of market relationships.

To better understand the rationale underlying the existing approaches to financial guarantees, interviews were conducted with representatives of the competent administrative authorities in CEFTA members.

4. Representatives of CEFTA members’ tourism authorities on travel guarantees

Interviews with representatives of government authorities and partnering institutions responsible for regulating and supervising the licensing of travel agencies and tour operators in CEFTA members were conducted between 8 and 16 June 2026. In BiH, in addition to state representatives, representatives of the entities and a number of cantons participated. The discussions identified several common issues concerning mandatory financial protection for tour operators.

Minimum value of guarantee. Although the minimum guarantee amount is not predefined by law in Albania, the insurers participating in the discussions argued that the legislation should establish a minimum amount of coverage. They suggested that the level of financial protection should be linked to turnover and differentiated

according to turnover categories, for example, up to EUR 1 million, up to EUR 5 million, and so forth. However, in Albania and other CEFTA members, data on tour operator turnover are available through the previous year's annual reports. Serbia is an exception, as its *eTurista* platform is updated daily: agencies must keep records of contracts, turnover, etc. in this system, including information on tourist travel guarantees. The problem is that access to *eTurista* is accessible only to certain members of the Serbian ministry in charge for tourism, so the information is not public. *eTurista* is connected with the Central Information System (CIS) also maintained by Serbian ministry in charge for tourism and with the Tourism register, maintained by Serbian Business Register in charge for licencing tour operators. Ministry in charge for tourism in Serbia is notified on daily basis if the financial guarantee of travel services is exceeded over prescribed minimum for the turnover reported in the application for licencing and request this guarantee to be increased. This makes the Serbian system comparatively sophisticated, but also considered more complicated and incomparable with systems applied in other CEFTA members or other European countries. Graded minimum financial guarantee linked to turnover is prescribed only in Serbia. In Montenegro and Republic of Srpska different guarantee amounts apply to tour operators and travel agents acting as resellers of tour operator packages. In Serbia and North Macedonia, no financial guarantee is prescribed for resellers i.e. travel agencies, meaning that principal responsibility for traveller protection rests with the tour operators. Representatives of the authorities interviewed indicated that history of frauds in both countries was one of the main reasons for establishing relatively high insurance requirements.

Types of risk to be covered. The interviews also demonstrated that the types of risks covered by financial guarantees vary significantly and that not all potential risks are necessarily covered by a single instrument. Consequently, different forms of partial insurance may be required. For example, Albanian insurers proposed coverage of up to EUR 35,000 in cases of injury or death, to be developed in cooperation with the competent ministry and the association of insurers. A financial bond covering professional indemnity could provide coverage of up to EUR 100,000. The interviews also demonstrated that the types of risks covered by financial guarantees vary significantly and that not all potential risks are necessarily covered by a single instrument. Consequently, different forms of partial insurance may be required. For example, Albanian insurers proposed coverage of up to EUR 35,000 in cases of injury or death, to be developed in cooperation with the competent ministry and the association of insurers. A financial bond covering professional indemnity could provide coverage of up to EUR 100,000. The potential liabilities are diverse and may include personal liability as well as liabilities arising from business relationships with hotels, transport providers, and other suppliers. The insurance requirements applicable to travel agencies and tour operators should therefore be clearly distinguished. Montenegro provides for a relatively broad scope of risks, including damages and failure to fulfil contractual obligations. North Macedonia covers persons and their belongings, while Albania provides coverage for liabilities arising from failure to fulfil obligations. By contrast, the system in the Federation of Bosnia

and Herzegovina focuses primarily on refunds and repatriation. Other jurisdictions additionally provide for accommodation and food expenses. These differences in the scope of coverage make the financial guarantees applied across CEFTA markets difficult to compare directly.

Recognition of the foreign insurance during temporary provision of services. CEFTA Joint Committee Decision No. 6/2024 enables the cross-border provision of services in other CEFTA party without establishment, for a limited period of time up to six months per year. However, the implementation of this Decision raised the questions concerning the recognition of financial guarantee issued in another CEFTA market. Next question concerns the situations in which the amount of foreign guarantee is lower than the amount required by the legislation of host country. This issue was raised in Albania, together with question concerning the conditions under which a guarantee can be triggered, as well as the question of its scope of coverage. According to Albanian insurers, these issues has to be addressed through legislative amendments, including provisions governing the rejection of prior declaration of intended cross border provision of services where, for example, the amount of the guarantee is insufficient or the relevant insurance coverage is excluded. One proposal was to establish local mandatory mechanism for handling claims arising in the host market, potentially through a dedicated committee or Bureau operating in a manner similar to the Green Card system. In Kosovo, the possibility of developing a new multilateral agreement among CEFTA parties was proposed, in order to operationalise the free movement of these services. In Montenegro, the discussion focused primarily on the requirement that a person temporarily providing cross-border tourism services must not have been previously convicted. Similar provisions exist in the legislation governing tourism activities of a branch manager of a tour operator in all other CEFTA parties.

An international insurance instrument was also proposed in Albania based on a model similar to the Green card system, as a possible multilateral insurance mechanism for travel agencies operating across the region. Such an arrangement could facilitate coverage in markets where insurance companies are unwilling to provide policies because they consider the insolvency risk of tour operators difficult to assess, as has reportedly occurred in Kosovo. In both Albania and Bosnia and Herzegovina, participants in the interview also noted that local issued financial guarantees do not necessarily cover risks arising from contracts concluded outside the domestic market. In BiH, on state and on Federal level, the legislation requires that insurance policy to be issued by a domestic insurance company. Although not always stated in tourism legislation, appears to apply in all other CEFTA members. In Kosovo, insurance companies have reportedly been unwilling to insure tour operators against insolvency under the new Law on tourism, arguing that the relevant risk cannot be reliably assessed. This has effectively halted the issuance of licences despite the fact that the bank guarantees are formally recognised as an instrument of financial protection. Under the current circumstances, established agencies enjoy a significant advantage because they can continue operating under the conditions that

applied before licensing became mandatory. Albanian insurers proposed the establishment of a standardized insurance certificate that would make the terms of coverage transparent and publicly available in advance, while ensuring that the amount of coverage is proportionate to the operator's turnover.

Conclusion

The mechanisms used to protect travellers against tour operator insolvency vary considerably across CEFTA markets and also differ from prevailing European approaches. Germany, for example, operates a mandatory state-supervised travel security fund financed through contribution from tour operators on per trip basis, while the Netherlands relies on private guarantee funds. Austria, meanwhile, prescribes a minimum level of insurance coverage for tour operators. In many other European countries, the legislation specifies the risks and obligations that must be covered, but leaves the assessment of the underlying business risk, the amount of coverage, and corresponding premium to financial institutions. Such assessments are generally based on the business and financial data of tourism service providers, including the volume of their recorded business activity.

In contrast, regulators of tour operators and travel agencies in CEFTA members generally do not base consumer protection requirements on the accounting data of business entities operating in their respective markets. Instead, they tend to prescribe fixed minimum amounts of financial security, implemented either through insurance policies or bank guarantees. Albania and the Federation of Bosnia and Herzegovina represent partial exceptions as their regulatory frameworks allow a greater role for insurers in assessing the underlying risk. However, insurers in Albania have also expressed a preference to statutory minimum amounts, arguing that they face difficulties in independently assessing business risks of tour operators. A similar position appears to prevail across the other CEFTA markets.

Why, then, is the assessment of tour operator's business risk so difficult in CEFTA markets?

One possible explanation is that these insurance markets remains relatively closed. Insurance companies that are not established or authorised in particular market are generally unable to compete for local business and offer their products to domestic tour operators. Greater openness of the insurance market, including the possibility of purchasing insurance from providers established in other CEFTA members could encourage the development of a more sophisticated risk-assessment models and increase competition among insurers.

A second possible explanation is the relatively high incidence of frauds in some CEFTA markets, a concern also raised by public officials during the interviews. This may be associated with inefficient judicial procedures and relatively mild sanctions, which may fail to provide sufficient deterrence to potential perpetrators. In such circumstances, accounting data concerning previously realised turnover may not

always provide a reliable basis for risk assessment. If a significant proportion of turnover is concealed and collected in cash, the recorded turnover will underestimate the actual volume of business activity and, consequently, the associated insolvency risk. Furthermore, experience in some markets indicates that businesses may establish new legal entities with little or no recorded turnover in the preceding year and consequently obtain lower insurance requirements or premiums. The bulk of the business activity may then be transferred to such entities, sometimes during the very year in which insolvency occurs and bankruptcy proceedings are initiated. Such practices further undermine the reliability of turnover-based risk assessment.

The overall conclusion is therefore that effective traveller protection against tour operator insolvency cannot be achieved solely by prescribing minimum financial guarantees. A more comprehensive approach requires significant improvements in the efficiency of bankruptcy proceedings and, more broadly, in the enforcement of regulations. Greater transparency of business activity, improved access to reliable financial data, more competitive insurance markets, and effective enforcement mechanisms would together provide a stronger basis for assessing business risk and determining appropriate levels of financial protection for travellers.

Ultimately, improving the efficiency of the insolvency and enforcement systems should be regarded as a prerequisite for the effective functioning of other market mechanisms, including the insurance and compensation systems designed to protect travellers against tour operator insolvency.

References

- Almansoori, A., Alghfeli, H., Nobanee, H. (2024). Travel Insurance Trends and Insights. SSRN Electronic Journal. <https://dx.doi.org/10.2139/ssrn.4978839>
- BGB. (2002). Bürgerliches Gesetzbuch in der Fassung der Bekanntmachung vom 2. Januar 2002 (BGBl. I S. 42, 2909; 2003 I S. 738), das zuletzt durch Artikel 6 des Gesetzes vom 23. Juli 2026 (BGBl. 2026 I Nr. 226) geändert worden ist.
- Charles G., (2019). Coordinating regulatory agencies for improved tourism policy in Tanzania, *Journal of Hospitality and Tourism Insights*, 2(1), 2-17. <https://doi.org/10.1108/JHTI-08-2018-0049>
- Chen, Z. (2021). The Tango Between Art.17(3) Brussels Ibis and Art.6(4)(b) Rome I under the Beat of Package Travel Directive. *Maastricht Journal of European and Comparative Law*, 28(6), 878-899. <https://doi.org/10.1177/1023263X211048595>
- Conway, L. (2019). Brexit: ATOL protection scheme, BRIEFING PAPER Number 8559, 5 November 2019, House of Commons Library, UK Parliament, <https://researchbriefings.files.parliament.uk/documents/CBP-8559/CBP-8559.pdf>

Council Directive 90/314/EEC of 13 June 1990 on package travel, package holidays and package tours (OJ L 158, 23.6.1990, p. 59, ELI: <http://data.europa.eu/eli/dir/1990/314/oj>)

Deacon C. (2025). Package travel reform and the UK: pioneering divergence in a closely aligned market or following in the shadow of the EU? In the: “Collective Commentary about the New Package Travel Directive“, Stewarts, (preprint)

DIRECTIVE (EU) 2015/2302 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, of 25 November 2015 on package travel and linked travel arrangements, amending Regulation (EC) No 2006/2004 and Directive 2011/83/EU of the European Parliament and of the Council and repealing Council Directive 90/314/EEC, Official Journal of the European Union, <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32015L2302>

DIRECTIVE (EU) 2026/1024 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, of 29 April 2026, Official Journal of the European Union, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202601024

Eisma, R.-L. V., Christie, P., & Hershman, M. (2005). Legal issues affecting sustainability of integrated coastal management in the Philippines. *Ocean & Coastal Management*, 48 (3-6), 336–359. <https://doi.org/10.1016/j.ocecoaman.2005.04.009>

Fact sheet | July 2018, Stronger EU protection for package, holidays https://commission.europa.eu/document/download/3ef0ed1e-13d9-4161-b4ff-f4bab38d3eb7_en?filename=factsheet_package_holidays_2018.06_en_web.pdf

Faure M.G. & Weber F. (2013). Security Mechanisms for Insolvencies in the Package Travel Sector: An Economic Analysis. *Journal of Consumer Policy*, 36 (4), 425–442. <https://doi.org/10.1007/s10603-013-9222-4>

FR Order No 2017-1717 on package travel and LTA, (2018). https://commission.europa.eu/document/download/a91d671eae1048efbbafaac14e488fe1_en?filename=fr_en_order_no_20171717_on_package_travel_and_lta.pdf

Gago A., Hruschka C. (2022). The persistence of legal uncertainty on EU citizens' access to social benefits in Germany. *Journal of European Social Policy*, 32 (5), 564–577. <https://doi.org/10.1177/09589287221095028>

Grant, D. (1996). The package travel regulations 1992: Damp squib or triumph of self-regulation? *Tourism Management*, 17(5), 319–321. [https://doi.org/10.1016/0261-5177\(96\)00031-3](https://doi.org/10.1016/0261-5177(96)00031-3)

Guvernul României - Ordonanță nr. 2/2018 din 02 august 2018, <https://anpc.ro/wp-content/uploads/2025/05/Ordonanta-nr.-2-2018-privind-pachetele-de-servicii-de-calatorie-si-serviciile-de-calatorie-asociate.pdf>

<https://drsf.reise/>, accessed 05.08.2026.

<https://www.boatbiketours.com/stichting-garantiefonds-reisgelden/>, accessed 05.08.2026.

European Commission. (2021). Revision of the Package Travel Directive: Combined evaluation roadmap/inception impact assessment (Ref. Ares (2021)5004927).

NL (2018). Act of 6 December 2017 amending Book 7 of the Civil Code, the Consumer Protection Enforcement Act and certain other acts in connection with the implementation of Directive (EU) 2015/2302 of the European Parliament and of the Council of 25 November 2015 on package travel and linked travel arrangements, amending Regulation (EC) No 2006/2004 and Directive 2011/83/EU of the European Parliament and of the Council and repealing Council Directive 90/314/EEC (OJ L 326, 11.12.2015, p. 1) ('the Act implementing the Directive on package travel and linked travel arrangements'), *Bulletin of Acts and Decrees of the Kingdom of the Netherlands* 2018 session https://commission.europa.eu/document/download/df6a8832deec4b3392caeb7469883ace_en?filename=nl_transposition_tra.pdf

Presidential Decree 7/2018. (2018). Government Gazette Of The Hellenic Republic, Issue No 12, https://commission.europa.eu/document/download/70c28850-0c5b-4b0e8e9394999b58b642_en?filename=el_en_presidential_decree_7_on_package_travel_and_lta.pdf

PRV (2018). Ordinance of the Federal Minister for Digitalization and Business Location on Package Holidays and Connected Travel Services (Package Travel Ordinance – PRV). StF: BGBl. II No. 260/2018 [CELEX No.: 32015L2302] <https://www.ris.bka.gv.at/geltendefassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20010321>

RSG - Reisesicherungsfondsgesetz. (2021). Travel Insurance Fund Act of 25. June 2021 (BGBl. I p. 2114), which is governed by Article 49 of the Act of 2. December 2024 (BGBl. 2024 I No 387) has been amended.

The Civil Aviation. (2012). (Air Travel Organisers' Licensing) Regulations 2012, UK Statutory Instruments 2012 No. 1017 <https://www.legislation.gov.uk/uksi/2012/1017/contents?utm>

The Package Travel and Linked Travel Arrangements Regulations 2018. (2018). UK Statutory Instruments, No. 634 <https://www.legislation.gov.uk/uksi/2018/634/part/5> www.sto-garant.nl, accessed 05.08.2026.

Zakon o pružanju usluga u turizmu. (2017). Narodne Novine 130/2017 (27.12.2017.) https://commission.europa.eu/document/download/0e197642-dd4c-4fd1-8d48-1be07b4dec0d_en?filename=hr-zakon_o_pruzanju_usluga_u_turizmu.pdf

Abiola Adeniyi
Kwame Nkrumah University of
Science and Technology,
Regentropfen University College
ORCID: 0000-0003-4773-0206

Kingsley Opoku Appiah
Kwame Nkrumah University of
Science and Technology
ORCID: 0000-0003-4915-1259

Beatrice Sarpong-Danquah
Kwame Nkrumah University of
Science and Technology
ORCID: 0000-0003-3195-5726

THE DIGITAL FRAUD DIAMOND: SPECIFYING THE DIGITAL DIMENSION OF CAPABILITY

UDK 343.72:004

DOI <https://doi.org/10.54318/czf.2026.aa.023>

Abstract

Digitalisation has changed not only the channels through which fraud is committed but also the capabilities required to exploit those channels. This paper revisits the Fraud Diamond and specifies capability along two related but distinguishable dimensions: human capability and digital capability. The study uses a systematic integrative review of 72 academic, professional and grey-literature sources identified through searches of major academic databases and authoritative practitioner sources. Of the 72 reviewed sources, 55 address at least one capability dimension, while 17 explicitly engage with digital capability; the latter comprise 15 digital-only studies and two studies addressing both human and digital capability. The remaining 38 studies address human capability without an explicit digital capability dimension, while 17 address neither capability dimension. Rather than presenting digital capability as a fifth element of the Fraud Diamond, the paper identifies a narrower theoretical gap: digital capability is visible in the literature but remains weakly integrated with the established capability construct and is rarely specified as an actor-level digital dimension of capability. Drawing on Whetten's criteria for theoretical contribution, the paper clarifies the construct domain, distinguishes digital capability from human capability and opportunity, and positions organisational digital preparedness as a contextual boundary condition. The paper treats capability as the first element of the Fraud Diamond to be digitally specified and identifies the digital specification of pressure, opportunity and rationalisation as a future research programme. The resulting DFD is therefore presented as a conceptual specification for empirical validation rather than a statistically validated model.

Key words: Digital Fraud Diamond; Digital Capability; Fraud Theory; Forensic Accounting; Digital Fraud; Fraud Risk.

JEL Classification: M41; M42; O33; K42.

1. Introduction

The history of fraud theory in accounting can be understood as a continuing response to the changing ways in which fraud has been committed. Cressey's (1953) study of individuals convicted of embezzlement provided one of the earliest systematic explanations of fraud through what became known as the Fraud Triangle. He argued that fraud commonly occurs when individuals experience pressure, identify an opportunity and rationalise their actions. The model subsequently became influential in auditing, internal control and fraud-risk assessment. Corporate scandals such as Equity Funding and ZZZZ Best further demonstrated the persistence of financial statement fraud and reinforced the need to understand the behavioural and organisational conditions that allow it to occur (Albrecht, Howe & Romney, 1984). These ideas were eventually reflected in professional auditing guidance, including SAS 82 (AICPA, 1997).

The large corporate failures of Enron, WorldCom and Parmalat at the beginning of the twenty-first century, however, showed that the Fraud Triangle did not fully explain how sophisticated fraud was perpetrated. In many of these cases, senior executives possessed the authority, knowledge and confidence necessary to override controls and manipulate financial reporting systems. Wolfe and Hermanson (2004) responded to this limitation by introducing capability as a fourth element, creating the Fraud Diamond. Their argument was straightforward: pressure, opportunity and rationalisation may create the conditions for fraud, but the perpetrator must also have the ability to exploit those conditions. This perspective has remained important in auditing and forensic accounting, particularly as financial statement fraud has become more complex (Rezaee, 2019). More recent research also shows that managerial communication, including the tone used in management discussion and analysis, can provide useful signals when assessing fraud risk (Rezaee, Pouryousof and Soleimani, 2024).

These theoretical developments were accompanied by stronger regulatory responses. The Sarbanes-Oxley Act of 2002 strengthened requirements for internal controls and executive accountability, while ISA 240 formally incorporated fraud-risk considerations into international auditing practice (IAASB, 2009). Over time, the Fraud Triangle and Fraud Diamond became familiar tools for auditors, forensic accountants, researchers and corporate governance practitioners.

The business environment has changed considerably since these models were developed. Organisations now operate through highly interconnected digital systems, and fraud increasingly takes place through technologies that traditional fraud theories were not designed to explain. Dorminey, Fleming, Kranacher, & Riley (2012), Free & Murphy (2015) and Free (2015) therefore question whether conventional behavioural models adequately capture the technological and systemic dimensions of contemporary fraud.

Recent events make this limitation particularly clear. The 2016 attack on The DAO, for example, involved the exploitation of a vulnerability in smart-contract code and resulted in the diversion of approximately USD 60 million (Meher, Shier, Giambattista, Gong, Fletcher, Sanayhie, Kim, & Laskowski, 2019). The case illustrates how fraud can be enabled by weaknesses in technology rather than simply by an insider exploiting conventional organisational controls. Deepfake technology presents an even newer challenge. In 2024, a finance employee was reportedly deceived into transferring USD 25 million after being led to believe that a deepfake video call involved the company's chief financial officer (Chen & Magramo, 2024). Evidence of the growing use of deepfakes in fraud further suggests that digital technologies are changing both the scale and nature of fraudulent activity (de Souza, 2026; Deloitte, 2024).

Blockchain provides another example. Although it was initially promoted as a technology capable of improving transparency and reducing fraud, its adoption has also introduced new governance, cybersecurity and auditability challenges. Many auditors still have limited knowledge of blockchain systems, particularly in smaller firms (Anis, 2023), while the decentralised nature of blockchain can challenge conventional approaches to assurance (Zhang, Ma, & Meng, 2025; Pimentel, Boulianne and Spence, 2025). Its architectural, scalability and cybersecurity limitations can themselves create opportunities for fraud (Bonsón and Bednárová, 2019). These developments suggest that contemporary fraud is no longer purely a behavioural problem; it is increasingly a socio-technical phenomenon arising from the interaction between people, technology, organisations and institutions.

It is against this background that this paper proposes the Digital Fraud Diamond (DFD) as a focused digital specification of the Fraud Diamond. The model retains the four established elements of pressure, opportunity, rationalisation and capability, and specifies capability as comprising two distinct dimensions, human capability and digital capability. Digital capability reflects the technology-mediated ability to understand, access, manipulate, automate and exploit digital technologies and infrastructures in ways that may facilitate fraud. The paper does not treat organisational digital preparedness as part of perpetrator capability; rather, it is positioned as a contextual boundary condition that can constrain the conversion of digital capability and digital opportunity into successful fraud.

Drawing on a systematic review of 72 studies, the paper develops propositions and initial indicators for the proposed capability specification. The Digital Fraud Diamond therefore contributes to the continuing development of fraud theory by making the digital dimension of perpetrator capability explicit while preserving the established meaning of capability in the Fraud Diamond. The paper also recognises that digitalisation affects all four Fraud Diamond elements and therefore identifies the digital specification of pressure, opportunity and rationalisation as a broader research direction.

2. Literature review and theoretical gap

2.1. Evolution of fraud theory and the extension literature

The study of fraud has evolved as scholars have sought to explain not only why people commit fraud, but also what enables them to carry it out successfully. Cressey's (1953) Fraud Triangle provides the traditional foundation, explaining fraud through the interaction of pressure, opportunity and rationalisation. Wolfe and Hermanson (2004) subsequently extended this framework by introducing capability, arguing that an individual must possess the characteristics, position, knowledge and confidence necessary to recognise and exploit an opportunity. Further extensions have added other behavioural and social dimensions. Crowe (2011), for instance, introduced arrogance through the Fraud Pentagon, while Vousinas (2019) developed the S.C.O.R.E. model, drawing attention to additional conditions such as ego and collusion. These developments demonstrate that fraud theory has continued to evolve in response to changing forms of fraudulent behaviour. A systematic review by Chimonaki, Papadakis and Lemonakis (2023) similarly confirms the substantial growth of research extending and applying fraud theories.

The increasing digitalisation of business and financial activities has created another important dimension of this evolution. Fraud can now be facilitated through artificial intelligence, blockchain, automated systems, online platforms and digital financial infrastructures, meaning that the ability to commit fraud may require more than conventional authority, intelligence or confidence. A perpetrator may also need to understand digital systems, manipulate data, configure technological tools, automate fraudulent activities or exploit technological vulnerabilities. Recent technology-oriented work reflects this development. For example, Zweers, Dey and Bhaumik (2025) proposed the AI-Fraud Diamond by introducing technical opacity, focusing on how the complexity or limited transparency of AI systems can conceal or facilitate harmful outcomes. However, technical opacity and digital capability represent different ideas. Technical opacity is primarily a system-level characteristic, whereas digital capability in the present study refers to the technology-mediated ability of the fraud-enabling actor. Similarly, digital capability should not be confused with opportunity: opportunity concerns the conditions that make fraud possible, such as weak controls or system vulnerabilities, while digital capability concerns the actor's ability to recognise, manipulate or exploit those conditions.

Consequently, the Digital Fraud Diamond (DFD) is positioned as a focused specification of the established Fraud Diamond rather than as a fifth-element replacement for existing fraud theories. The model retains pressure, opportunity, rationalisation and capability, and divides capability into a human and a digital dimension. The central argument is that conventional human capability may not fully capture the specialised technology-mediated abilities required to exploit increasingly digital environments. The contribution of the DFD is therefore to distinguish the perpetrator's broader human and positional capability from the specialised digital capability through which digital systems, data and assets may be manipulated,

exploited or automated. This refinement remains consistent with the broader evolution of fraud theory while avoiding the implication that digital capability constitutes an additional vertex beside capability.

2.2. Digitalisation and the Metamorphosis of fraud

Digitalisation is changing not only how fraud is committed but also the conditions that can lead to fraud. It can affect all four elements of the Fraud Diamond. Pressure may increase when employees are constantly monitored through digital systems, assessed against algorithm-driven targets or exposed to greater online visibility. Opportunity may arise when organisations have weak access controls, poorly protected applications, inadequate system monitoring or vulnerabilities in technologies such as smart contracts. Rationalisation may also change in digital environments. For example, people involved in fraudulent activities may convince themselves that they are not fully responsible because the action was carried out through a computer program, an algorithm or a decentralised platform. In such situations, the technology may be seen as responsible for the outcome rather than the individual who deliberately used it (De Filippi and Hassan, 2021). Human capability remains important because fraud still requires qualities such as confidence, intelligence, authority, persistence and the ability to bypass or influence controls. However, digital environments introduce another important question: does the person have the technical knowledge and skills needed to use the digital system effectively? This is where digital capability becomes important.

Taken together, these observations indicate that digitalisation bears on each of the four elements of the Fraud Diamond, and that a complete digital specification of the model would require the digital aspect of pressure, opportunity, rationalisation and capability each to be developed as a construct in its own right. That is the larger programme to which the Digital Fraud Diamond points. The present paper takes capability as its first case, because it is there that the review reported in section 2.3 finds the literature divided. The digital specification of the remaining three elements is set out as a research direction below.

2.3. Evidence from the 72-Study review

The 72-study review provides useful evidence on how capability has been treated in the fraud literature. The studies were classified according to whether they explicitly addressed human capability, digital capability, or both. The classification produced four groups in Table 1: studies addressing neither dimension (17), studies addressing human capability only (38), studies addressing digital capability only (15), and studies addressing both human and digital capability (2). Together, the findings show that capability remains an important element in fraud research, but the attention given to its human and digital dimensions is uneven.

Table 1: Classification of the 72 reviewed studies by human and digital capability

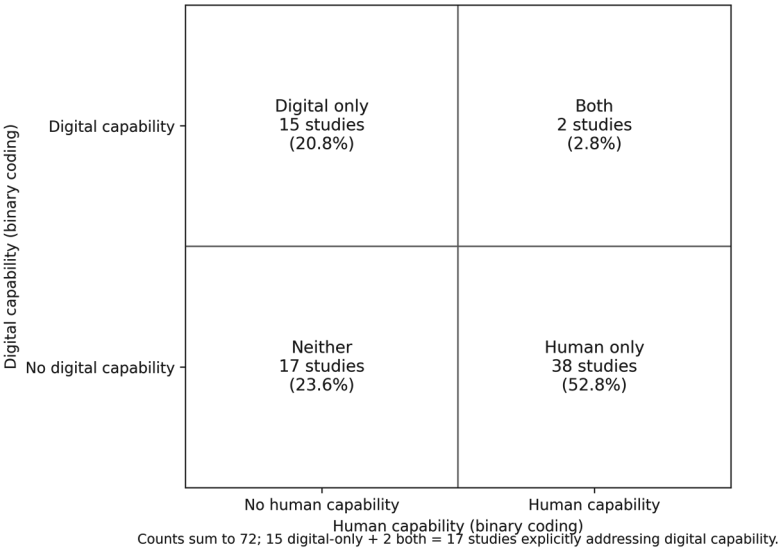
Human capability	Digital capability	Studies	Percentage
No	No	17	23.6%
Yes	No	38	52.8%
No	Yes	15	20.8%
Yes	Yes	2	2.8%
Total		72	100.0%

Source: Authors' Own Research

The distribution reveals three important patterns. First, human capability remains the dominant focus, with 38 studies (52.8%) explicitly addressing it without a corresponding digital capability dimension. This reflects the strong influence of traditional fraud theories, particularly the Fraud Diamond, where capability is largely understood in terms of characteristics such as authority, intelligence, confidence, position and the ability to overcome or influence controls. Second, digital capability is beginning to receive attention, appearing in 17 studies (23.6%), comprising 15 studies that focus on digital capability and two that consider both human and digital capability. This suggests that the growing use of digital technologies in business and financial activities is creating increasing interest in the technological abilities required to facilitate or exploit fraud.

Third, and most importantly, only two of the 72 studies (2.8%) bring human and digital capability together within the same analysis. This limited overlap is theoretically important. It suggests that the literature has generally considered the capability of fraud perpetrators from either a conventional human perspective or a technology-related perspective, rather than examining how these two forms of capability may coexist and complement each other. The evidence therefore points to a need for a framework that preserves the established understanding of human capability even though giving explicit attention to the additional technical abilities that may be required to commit fraud in increasingly digital environments. It is this gap in the integration of the two dimensions that provides the empirical basis for introducing digital capability as a distinct dimension of capability of the Digital Fraud Diamond.

Figure 1: Distribution of the 72 Reviewed Studies by Human and Digital Capability



Source: Authors’ review and coding of the 72 studies.

3. Methodology

3.1. Review design and Search Strategy

This study adopts a systematic integrative review to examine how fraud theory has developed from its traditional foundations to more recent explanations of fraud in digital environments. An integrative review is appropriate because the study brings together different strands of literature, including foundational fraud theories, subsequent extensions of these theories, and emerging research on technology-enabled fraud. The review therefore considers fraud research from accounting, auditing, information systems and cybersecurity perspectives. Foundational works were traced from 1953, beginning with Cressey’s Fraud Triangle, while contemporary studies published up to 2025 were considered to capture recent developments in digital fraud (See Table 2). This approach allows the review to connect the historical development of fraud theory with the changing technological environment in which fraud now occurs.

Table 2: Screening and Eligibility Criteria

Stage	Description of Action	Records Remaining	Records Excluded
Identification	Records retrieved from academic databases, practitioner sources, and grey literature (Scopus, Web of Science, Emerald, SSRN, AICPA, Chainalysis, Big Four, etc.)	562	—
Duplication	Duplicate records removed	498	64
Title and Abstract Screening	Titles and abstracts reviewed; irrelevant technical or non-fraud papers excluded	144	354
Full Text Eligibility	Theoretical, fraud-related or digital-fraud relevance assessed	102	42
Inclusion	Studies included in final synthesis based on quality, geography, and balance between foundational and contemporary works	72	30

Source: Authors' Compilation

The literature search was conducted across major academic databases and publisher platforms, including Scopus, Web of Science, Emerald Insight, ScienceDirect, Wiley Online Library, Taylor & Francis and SpringerLink. To complement the academic literature, relevant professional and selected grey-literature sources were also considered, including materials from ACFE, AICPA, CIMA and ISACA (Smith, 2024), as well as selected reports from recognised cybersecurity and financial-crime organisations and SSRN. The search used combinations of terms relating to fraud theory, capability and digital technologies. Key search expressions included “fraud diamond AND blockchain,” “fraud diamond AND digital capability,” “fraud triangle AND AI,” “fraud theory AND digital fraud.” Foundational studies, particularly Cressey (1953) and Wolfe and Hermanson (2004), were retained because they provide the theoretical foundation for understanding the development of fraud theory and the later emergence of capability as a central element.

3.2. Coding Procedure

Following the screening process, data from the 72 studies that were retained were organised in an extraction register containing the main bibliographic and substantive details of each study, including the author, year, title, geographical context, study type, objectives, findings, identified gaps and relevance to the Digital Fraud Diamond. The information recorded for each study was then used to determine whether the study explicitly addressed human capability and/or digital capability. This provided a consistent basis for moving from the broader literature review to a more focused assessment of how capability has been treated in existing fraud research.

The coding was conducted by the researcher and checked by an assisting investigator. Both reviewers assessed whether each study explicitly addressed human capability, digital capability, both, or neither, using the predefined coding definitions. Borderline cases were jointly reviewed and resolved through discussion, with reference to the operational definitions and the substantive content of the study. This process was used to improve the consistency and credibility of the final classification.

Each study was subsequently coded on two binary dimensions: human capability and digital capability. Human capability was coded 1 where the study explicitly discussed the perpetrator's capability, competence, authority, positional power, confidence, intelligence or related characteristics within a fraud context, and 0 where these aspects were not explicitly addressed. Digital capability was coded 1 where the study explicitly connected fraud with technology-related skills, technical knowledge, digital access, technological proficiency or the ability to operate, manipulate or exploit digital systems, and 0 where such a dimension was not explicitly addressed. Importantly, the mere mention of a digital technology did not automatically result in a digital-capability code. For example, a study could discuss blockchain, artificial intelligence or cryptocurrency as part of the fraud environment without necessarily examining the perpetrator's ability to use or exploit that technology.

The two binary codes were then combined to produce a 2×2 classification of the 72 studies. This resulted in four groups: 38 studies addressing human capability only, 15 addressing digital capability only, two addressing both human and digital capability, and 17 addressing neither dimension. Thus, 55 studies addressed at least one of the two capability dimensions, while 17 explicitly addressed digital capability when the two dimensions are considered together (15 digital-only + 2 both). This classification provides the basis for the analysis presented in Table 1 and Figure 1 and, more importantly, allows the review to move beyond simply counting studies towards examining how capability has been conceptualised in the existing literature.

3.3. Analytical Approach

The analysis combined qualitative thematic synthesis and structured binary classification to examine how capability has been treated in the existing fraud literature. The thematic synthesis was guided by the principle that evidence from multiple studies can be systematically coded, compared and brought together to identify recurring concepts and broader analytical themes (Thomas and Harden, 2008). In line with this approach, the substantive information extracted from the 72 retained studies was examined to identify recurring patterns concerning the meaning of capability, the role of digital technologies in creating or changing opportunities for fraud, and the extent to which technology was treated as a distinct theoretical condition rather than simply as part of the fraud environment. This approach is also consistent with Braun and Clarke's (2006) view of thematic analysis as a flexible method for identifying and interpreting meaningful patterns within qualitative material. The thematic analysis therefore helped the study move beyond simply

counting studies and provided the conceptual basis for understanding how human and digital capability have been treated across the literature.

The purpose of the classification was not to establish a causal relationship between human and digital capability, but to identify patterns in how the two dimensions have been conceptualised and examined in the literature. The results show that 38 studies addressed human capability only, 15 addressed digital capability only, two addressed both dimensions, and 17 addressed neither. Thus, while digital capability is increasingly visible in the literature, its integration with conventional human capability remains limited, with only 2 of the 72 studies (2.8%) explicitly addressing both dimensions. This pattern provides an important empirical basis for the Digital Fraud Diamond: rather than suggesting that human and digital capability are unrelated, it indicates that the existing literature has largely treated them as separate areas of inquiry. The DFD therefore seeks to bring these dimensions into a common theoretical framework, while maintaining a clear distinction between what the perpetrator is capable of doing through conventional human attributes and what the perpetrator is capable of doing through technology-mediated skills and knowledge.

4. The Digital Fraud Diamond: Construct Specification and the Theoretical Contribution

4.1. Conceptualisation of Digital Capability

Digital capability is defined as the technology-mediated ability of a fraud-enabling actor to understand, access, configure, manipulate, automate or exploit digital systems, data or digital assets in ways that facilitate fraudulent action. The construct focuses specifically on the actor's ability to use technology rather than on the sophistication of the technology itself. Organisational digital preparedness, including monitoring systems, forensic analytics, staff training and incident-response capacity, is treated as part of the broader environment within which digital capability operates rather than as part of the construct.

This definition responds directly to the construct-development test proposed by Whetten (1989), which requires theoretical contributions to clarify what is being added and how it differs from existing concepts. Digital capability therefore goes beyond general technical knowledge or the mere use of digital technology. Its relevance arises when an actor can apply technology-related knowledge and skills to execute, adapt, conceal or scale fraudulent behaviour. This distinction gives digital capability a specific role within the Digital Fraud Diamond and separates it from opportunity, which concerns the conditions that make fraudulent action possible.

4.2. Digital Capability and Human Capability

Wolfe and Hermanson (2004) position capability as the element that enables a person to recognise an opportunity and convert it into fraudulent action. In their Fraud Diamond, capability is associated with factors such as the individual's position or

function, intelligence and creativity, confidence or ego, and the ability to coerce others. Subsequent applications of the Fraud Diamond have similarly treated capability as a perpetrator-level characteristic that determines whether an individual is able to exploit an available opportunity (Wolfe and Hermanson, 2004; Dorminey et al., 2012). Human capability in this study therefore captures the general personal and positional capacity of the perpetrator, rather than technical competence alone (Table 3).

Table 3: Conceptual Distinction vs Human Capability

Dimension	Human capability	Digital capability
Core meaning	Ability to exploit fraud conditions through authority, confidence, intelligence, positional access and related personal characteristics.	Ability to operate through digital systems using technology-mediated knowledge, access, proficiency and adaptability.
Primary locus	Actor/person.	Actor/person, but expressed through digital technologies.
Illustrative mechanism	Override controls, persuade others, exploit authority or coordinate complex fraud.	Manipulate digital environments, exploit software or data weaknesses, use synthetic media or operate digital assets.
Why distinct?	Explains who has the personal and positional capacity to carry out the act.	Explains the additional technical means needed to execute or scale fraud in digital environments.

Source: Authors' Own Table

Digital capability extends this logic into technology-mediated environments. It concerns what the perpetrator is able to do through digital systems, including the ability to understand technological processes, manipulate digital information, exploit system vulnerabilities, automate activities or operate digital assets. The distinction is therefore not between a human actor and a technological actor; both constructs remain located in the perpetrator. Rather, human capability captures the broader personal and positional capacity to carry out fraud, while digital capability captures the specialised technological capacity through which that fraud can be executed in a digital environment.

The two capabilities can consequently coexist and reinforce one another. An individual may possess considerable authority, confidence and intelligence but have limited digital proficiency, while another may possess substantial digital expertise without occupying a powerful organisational position. A sophisticated fraud may require both: human capability may enable the actor to obtain access, influence others or override controls, while digital capability may enable the actor to

manipulate the underlying technology. Treating digital capability as a distinct dimension therefore preserves the established meaning of capability in the Fraud Diamond while recognising an additional form of capability that becomes increasingly important as fraudulent activities become technologically mediated.

4.3. Digital Capability vs Opportunity

Opportunity refers to the external or situational conditions that create an opening for fraud, such as weak internal controls, inadequate monitoring, poor segregation of duties or vulnerabilities within a digital system. Digital capability, by contrast, resides with the actor. For example, a vulnerability in a smart contract may create an opportunity for fraud, but recognising the vulnerability, understanding how it works and knowing how to manipulate and exploit it require digital capability. The existence of an opportunity, therefore, does not automatically mean that an individual has the ability to exploit it. Likewise, an actor may possess strong digital capability but still be unable to commit fraud where effective controls prevent access or exploitation.

This distinction is particularly relevant in digital environments, where technological weaknesses can create opportunities that are only meaningful to actors with the appropriate technical knowledge and skills. Research on smart-contract security, for example, shows that vulnerabilities can provide exploitable openings, while identifying, understanding and testing those vulnerabilities requires specialised technical knowledge (Qian et al., 2022; Garanina, Klamm and Street, 2022; Smith, 2024). Digital capability should therefore be understood as the actor's ability to exploit a digital opportunity, rather than simply another way of describing the opportunity itself.

4.4. Positioning the DFD against Existing Extension

The DFD is not presented as the first extension of fraud theory beyond the Fraud Diamond. The Fraud Pentagon introduced arrogance (Crowe, 2011), the Fraud Hexagon incorporated collusion through the S.C.O.R.E. framework (Vousinas, 2019), and recent scholarship has proposed technical opacity as a technology-specific condition in AI-enabled fraud (Zweers, Dey and Bhaumik, 2025). Other studies apply existing fraud extensions to cybercrime and digital fraud without necessarily introducing a new construct (Chimonaki, Papadakis and Lemonakis, 2023).

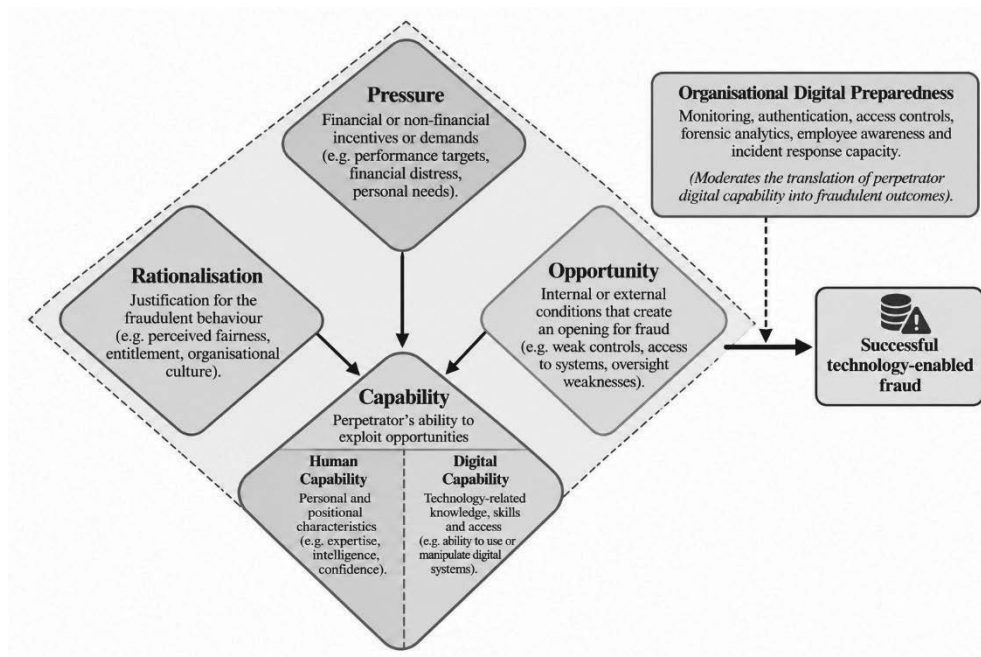
The DFD addresses this issue by distinguishing digital capability from both general human capability and digital opportunity. Human capability concerns the perpetrator's broader personal and positional ability to recognise, obtain and exploit an opportunity, whereas digital capability concerns the specialised ability to understand, manipulate, automate or exploit digital systems, data and digital assets. Digital opportunity, in contrast, refers to vulnerabilities or weaknesses within the environment that create an opening for fraud. These dimensions may coexist and

reinforce one another, but they perform different explanatory roles. A technological vulnerability may create an opportunity, for example, without every potential perpetrator possessing the technical ability to exploit it. Similarly, a highly capable digital actor may be unable to commit fraud where effective controls prevent access. The contribution of the DFD therefore lies not simply in adding “technology” to the Fraud Diamond, but in making the perpetrator's technology-mediated capability theoretically visible as a distinct analytical dimension.

4.5. DFD Propositions

The Digital Fraud Diamond (DFD) extends the Fraud Diamond by specifying capability along two dimensions: human capability and digital capability. Human capability captures the broader personal and positional characteristics that enable an actor to recognise and exploit an opportunity, while digital capability captures the technology-mediated ability required to understand, manipulate, automate or exploit digital systems. The model therefore retains the four conditions of the Fraud Diamond pressure, opportunity, rationalisation and capability while specifying the capability condition in human and digital terms. Importantly, opportunity creates an opening for fraud, whereas digital capability concerns the actor's ability to exploit that opening.

Figure 2: Conceptual Specification of the Digital Fraud Diamond



Source: Authors' conceptualisation based on Cressey (1953), Wolfe and Hermanson (2004), Whetten (1989), and the reviewed digital-fraud literature

Organisational digital preparedness is positioned outside the four Fraud Diamond conditions as a moderating boundary condition. Strong monitoring, authentication, access controls, forensic analytics and response mechanisms can constrain the ability of digitally capable perpetrators to convert digital opportunities into successful fraud. Figure 2 therefore presents the four conditions of the Fraud Diamond, with capability specified along its human and digital dimensions, and with organisational digital preparedness constraining the translation of digital capability into successful technology-enabled fraud. Consequently, the study proposes the following:

P1: Digital capability provides additional explanatory value where specialised technological skills are required to exploit a fraud opportunity.

P2: Technology-enabled fraud is most likely where pressure, opportunity and rationalisation are present and capability is present in both its human and its digital dimension.

P3: Digital capability is distinct from opportunity because it reflects the actor's ability to exploit an opportunity, whereas opportunity represents the condition that makes fraud possible.

P4: Organisational digital preparedness weakens the conversion of digital capability into successful fraud by strengthening monitoring, authentication, controls, forensic analytics and response capacity.

4.6. Development of Construct and Proposed Measurement

As a conceptual study, this paper develops an initial measurement framework for digital capability to support its operationalisation in future empirical research. The proposed indicators are grounded in the reviewed literature and focus on the perpetrator's ability to understand, use, manipulate and adapt to digital technologies in ways that may facilitate fraud. The construct is deliberately distinguished from organisational digital preparedness, which concerns the organisation's defensive and monitoring capabilities. Appendix 1 therefore separates the proposed digital-capability indicators from organisational digital preparedness, which is treated as the boundary condition rather than as part of the core construct.

The framework captures five related dimensions of digital capability: technical proficiency, AI-enabled capability, digital-asset capability, exploitation capability and adaptive digital capability. Together, these dimensions reflect the central argument of the DFD that digital capability goes beyond mere access to technology; it concerns the ability of the fraud-enabling actor to effectively operate within and exploit digital environments. The framework therefore provides a theoretically grounded basis for future empirical operationalisation and further refinement of the construct.

5. Discussion and Theoretical Contribution

The findings of the 72-study review suggest that the main gap is not an absence of research on digital capability, but its limited integration with the established understanding of fraud capability. Seventeen studies explicitly addressed digital capability, while 38 focused on human capability without an explicit digital dimension and only two considered both. This pattern reflects the broader evolution of fraud theory, from the Fraud Triangle to the Fraud Diamond and subsequent extensions (Dorminey et al., 2012; Chimonaki, Papadakis and Lemonakis, 2023). Wolfe and Hermanson (2004) introduced capability to explain why some individuals can turn an opportunity into fraudulent action. The DFD builds on this logic by arguing that, in increasingly digital environments, perpetrators may require specialised technological knowledge and skills in addition to conventional personal and positional capabilities. Digital capability is therefore presented as a distinct but complementary dimension of perpetrator capability, rather than as a replacement for human capability. The model also clarifies the distinction between digital capability and opportunity: opportunity represents the condition that creates an opening for fraud, whereas digital capability represents the actor's ability to recognise and exploit that opening.

The DFD consequently makes three related contributions to fraud theory. First, it provides clearer construct boundaries between human capability, digital capability and opportunity, reducing the risk of treating technological vulnerabilities and perpetrator abilities as the same phenomenon. Second, it offers a more socio-technical understanding of fraud, recognising that contemporary fraudulent behaviour increasingly emerges from the interaction between human actors and digital environments, complementing later extensions of fraud theory such as the S.C.O.R.E. model (Vousinas, 2019). Third, it positions organisational digital preparedness as a contextual condition rather than as another element of perpetrator capability. Strong monitoring, authentication, access controls, forensic analytics and incident-response mechanisms may constrain the extent to which a digitally capable actor can convert an opportunity into successful fraud. This also has practical implications for auditors and forensic accountants, who may need to consider technological capabilities alongside conventional fraud-risk indicators, particularly as contemporary auditing standards place renewed emphasis on fraud risk assessment and response (IAASB, 2026). Overall, the contribution of the DFD is therefore not to claim that digital fraud or technology has been absent from previous research, but to provide a more focused integration of human capability, digital capability and digital opportunity within the established Fraud Diamond framework.

6. Limitations

This study is conceptual in design and does not include primary data collection to empirically validate the proposed construct of digital capability. To counter this limitation, the indicators were derived from a wide body of empirical and practitioner

research to ensure that the framework is anchored in established evidence rather than speculation. Another limitation lies in the reliance on secondary sources, which may introduce bias or gaps in coverage. This risk was mitigated by systematically reviewing prior literature and aligning the indicators with validated constructs in fraud and auditing studies. Finally, the scope of the study is restricted to theory development rather than empirical testing across contexts. To address this, the paper sets out clear propositions, measurement items, and sampling strategies that provide a pathway for future research to test, refine, and validate the Digital Fraud Diamond in diverse organisational environments. A further limitation is that no formal inter-coder reliability statistic was calculated for the binary literature coding. The study-level register and coding rules are nevertheless supplied in Appendix 2 to make the classification transparent and reproducible.

7. Future research agenda

1. The digital dimensions of pressure, opportunity and rationalisation should be specified on the pattern applied here to capability, so that the Digital Fraud Diamond can be completed as a model rather than extended at a single element.
2. Validate the proposed digital-capability construct and refine the indicators through exploratory and confirmatory empirical studies across highly digitised organisations.
3. Examine whether digital capability strengthens or conditions the relationship between human capability and technology-enabled fraud, particularly where perpetrators possess different combinations of authority and technical expertise.
4. Test the explanatory value of digital capability across industries and jurisdictions, including whether governance quality, regulatory strength and technological maturity alter its relationship with fraud outcomes.
5. Examine digital capability and fraud risk in blockchain-based organisations, cryptocurrency exchanges and decentralised-finance environments, with particular attention to the interaction between technological vulnerabilities and perpetrator capability.
6. Investigate how auditors and forensic accountants assess fraud risk when evidence is generated or manipulated using artificial intelligence, synthetic media and other emerging technologies.
7. Compare the digital capabilities of audit and forensic-accounting firms and examine whether differences in digital skills, forensic infrastructure and analytical capacity affect fraud detection, audit quality and stakeholder confidence.
8. Examine how individual-level digital capability and organisational digital preparedness interact in multinational organisations exposed to cross-border, technologically mediated fraud.

Conclusion

This paper revisits the Fraud Diamond in response to the growing technological mediation of fraud and proposes the Digital Fraud Diamond as a specification of the digital dimension of perpetrator capability. The 72-study review shows that capability is widely recognised in fraud research, but human and digital treatments remain largely separated: 38 studies address human capability only, 15 address digital capability only, two address both, and 17 address neither. The evidence therefore supports a narrower theoretical gap concerning the limited integration of human and digital capability rather than an absence of research on digital capability.

The proposed DFD retains pressure, opportunity, rationalisation and capability as the four established conditions of the Fraud Diamond, while specifying capability along human and digital dimensions. Digital capability is defined at the actor level and distinguished from opportunity, which represents the external condition that creates an opening for fraud. Organisational digital preparedness is positioned as a contextual boundary condition that may constrain the conversion of digital capability and opportunity into successful technology-enabled fraud. The paper also makes clear that digitalisation affects all four Fraud Diamond conditions; the present study focuses on capability because this is where the literature review identifies the clearest division between human and digital treatments.

The contribution is therefore a focused conceptual refinement rather than a claim of unprecedented extension. The proposed construct, propositions and initial indicators provide a basis for empirical validation, while the digital specification of pressure, opportunity and rationalisation remains an important research direction. For auditors, forensic accountants and regulators, the framework highlights the need to consider not only the behavioural and organisational conditions surrounding fraud but also the specialised digital abilities that may allow perpetrators to exploit technology-enabled opportunities.

References

- Agenda Item, D. (2009). IAASB CAG PAPER.
- AICPA (1997), "The CPA Letter", January/ February, pp. 1-14.
- Albrecht, S., Howe, K., & Romney, M. (1984). *Deterring fraud: The internal auditor's perspective*. Institute of Internal Auditors Research Foundation.
- Anis, A. (2023). Blockchain in accounting and auditing: unveiling challenges and unleashing opportunities for digital transformation in Egypt. *Journal of Humanities and Applied Social Sciences*, 5(4), 359-380.
- Azimi, A. (2025). *Digital Platform Durability: Examining the Impact of Structural Attributes* (Doctoral dissertation, RMIT University).

- Bonsón, E., & Bednárová, M. (2019). Blockchain and its implications for accounting and auditing. *Meditari Accountancy Research*, 27(5), 725–740. <https://doi.org/10.1108/MEDAR-11-2018-0406>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Chainalysis. (2024). *The 2024 geography of cryptocurrency report*. Chainalysis Research.
- Chen, H., & Magramo, K. (2024, February 4). Finance worker pays out \$25 million after video call with deepfake “chief financial officer.” *CNN*.
- Chimonaki, C., Papadakis, S., & Lemonakis, C. (2023). Perspectives in fraud theories - A systematic review approach. *F1000Research*, 12, 933. <https://doi.org/10.12688/f1000research.131896.1>
- Cressey, D. R. (1953). *Other people's money: A study in the social psychology of embezzlement*. Free Press.
- Crowe, H. (2011). *Putting the Freud in fraud: Why the fraud triangle is no longer enough*. Horwath, Crowe LLP.
- De Filippi, P., & Hassan, S. (2021). Smart contracts. *Internet Policy Review*, 10(2).
- de Souza, R. R. M. (2026). Legal remedies and regulatory frameworks to combat ai-driven deepfakes. In *Mitigating the Risks of AI Deepfakes* (pp. 94-116). CRC Press.
- Deloitte. (2024). *Deepfake banking fraud risk on the rise*. Deloitte Insights.
- Dorminey, J., Fleming, A. S., Kranacher, M. J., & Riley, R. A. (2012). The evolution of fraud theory. *Issues in Accounting Education*, 27(2), 555–579.
- Financial Sector Conduct Authority. (2025). *FSCA regulatory actions report: 1 April 2024–31 March 2025*. <https://www.fsca.co.za/>
- Financial Sector Conduct Authority. (2024). *FSCA integrated report 2023/2024*. Financial Sector Conduct Authority.
- Free, C. (2015). Looking through the fraud triangle: A review and call for new directions. *Meditari Accountancy Research*, 23(2), 175-196.
- Free, C., & Murphy, P. R. (2015). The ties that bind: The decision to co-offend in fraud. *Contemporary Accounting Research*, 32(1), 18-54.
- Garanina, T., Klamm, B., & Street, D. (2022). Blockchain and auditing: Implications for smart contract assurance. *Journal of International Accounting Research*, 21(3), 75–98.
- Global Initiative Against Transnational Organized Crime. (2022). *Technology and organised crime: The role of blockchain in fraud and financial crime*. Global Initiative Against Transnational Organized Crime.

- International Auditing and Assurance Standards Board. (2025). *ISA 240 (revised): The auditor's responsibilities relating to fraud in an audit of financial statements*. IAASB.
- International Auditing and Assurance Standards Board. (2026). *Understanding the International Standard on Sustainability Assurance 5000*. <https://www.iaasb.org/focusareas/understanding-international-standard-sustainability-assurance-5000>.
- Maras, M. H., & Alexandrou, A. (2019). Determining authenticity of video evidence in the age of artificial intelligence and in the wake of deepfake videos. *The International Journal of Evidence & Proof*, 23(3), 255–262.
- Mehar, M. I., Shier, C. L., Giambattista, A., Gong, E., Fletcher, G., Sanayhie, R., ... & Laskowski, M. (2019). Understanding a revolutionary and flawed grand experiment in blockchain: The DAO attack. *Journal of Cases on Information Technology (JCIT)*, 21(1), 19–32.
- Organized Crime and Corruption Reporting Project. (2025, January 2). *Behind the decision (Indonesia): How OCCRP's "Person of the Year" highlights the fight against corruption*. <https://www.occrp.org/en/announcement/behind-the-decision-indonesia-how-occrps-person-of-the-year-highlights-the-fight-against-corruption>.
- Pimentel, E., Boulianne, E., & Spence, C. (2025). When audit confronts blockchain. *Accounting, Auditing & Accountability Journal*, 38(4), 1149–1173. <https://doi.org/10.1108/AAAJ-12-2023-6768>
- Qian, W., Li, J., Wang, T., & Chen, H. (2022). Detecting vulnerabilities in Ethereum smart contracts: A systematic literature review. *Future Generation Computer Systems*, 131, 1–15.
- Raju, B., Vinayakumar, R., & Poornachandran, P. (2025). An elegant intellectual engine towards automation of smart contract vulnerability detection. *Scientific Reports*, 15, Article 14567.
- Rezaee, Z. (2019). *Business sustainability, corporate governance, and organizational ethics*. John Wiley & Sons.
- Rezaee, Z., Pouryousof, A., & Soleimani, A. (2024). An examination of management tone in the MD&A in the context of fraud pentagon. *Journal of Forensic Accounting Research*, 9(1), 167–193.
- See IAASB Proposed Strategy for 2020–2023 and Work Plan for 2020–2021. Available at: <https://www.iaasb.org/publications/proposed-strategy-2020-2023-and-work-plan-2020-2021>.
- Smith, S. Z. (2024). How to audit smart contracts. *ISACA Journal*.

South African Banking Risk Information Centre. (2024). *Annual crime statistics 2023*. SABRIC. <https://www.sabric.co.za/wp-content/uploads/2024/11/SABRIC-Annual-Crime-Stats-2023.pdf>.

Teneva, I. (2025, August 6). *AI in accounting and finance: Navigating rapid change*. AICPA & CIMA. <https://www.aicpa-cima.com/professional-insights/article/ai-in-accounting-and-finance-navigating-rapid-change>.

Thomas, J., & Harden, A. (2008). Methods for the thematic synthesis of qualitative research in systematic reviews. *BMC Medical Research Methodology*, 8, Article 45.

Vidal, F. R., Ivaki, N., & Laranjeiro, N. (2024). Vulnerability detection techniques for smart contracts: A systematic literature review. *Journal of Systems and Software*, 217, 112160.

Vousinas, G. L. (2019). Advancing theory of fraud: The S.C.O.R.E. model. *Journal of Financial Crime*, 26(1), 372–381. <https://doi.org/10.1108/JFC-12-2017-0128>

Whetten, D. A. (1989). What constitutes a theoretical contribution? *Academy of Management Review*, 14(4), 490–495.

Wolfe, D. T., & Hermanson, D. R. (2004). The fraud diamond: Considering the four elements of fraud. *The CPA Journal*, 74(12), 38–42.

Zhang, Y., Ma, Z., & Meng, J. (2025). Auditing in the blockchain: A literature review. *Frontiers in Blockchain*, 8, Article 1549729.

Zweers, B., Dey, D., & Bhaumik, D. (2025). The AI-Fraud Diamond: A novel lens for auditing algorithmic deception. *arXiv*. <https://arxiv.org/abs/2508.13984>

Appendix 1

Proposed Indicators for the Digital Capability Construct

Part A: Actor-Level Digital Capability

Proposed Indicators	Source(s)
I have the ability to use advanced digital tools such as coding languages, encryption tools, or penetration testing software.	Qian et al. (2022); Raju et al. (2025)
I have access to generative artificial intelligence tools (e.g. deepfake applications, voice cloning platforms).	Maras and Alexandrou (2019); Chen and Magramo (2024)
I am knowledgeable about how to create, store, and transfer digital assets such as cryptocurrencies.	Chainalysis (2024)
I can manipulate or bypass digital platforms through phishing scripts, malware, or other cyber exploitation tools.	Vidal, Ivaki, & Laranjeiro, (2024); Azimi (2025)
I regularly use digital anonymity tools such as VPNs or TOR networks when accessing online systems.	Organized Crime and Corruption Reporting Project [OCCRP], (2025)

Part B: Organisational Digital Preparedness (Boundary Condition)

Proposed Indicators	Source(s)
Our organisation employs forensic analytics tools (e.g. AI-driven anomaly detection, blockchain auditing, real-time transaction monitoring).	Teneva (2025)
Our organisation has identity verification systems that can detect manipulated media such as deepfakes (e.g. biometric verification, multi-factor authentication).	Financial Sector Conduct Authority [FSCA], (2025)
Our organisation has a clearly defined cyber incident response process, including reporting channels and dedicated fraud response teams.	SABRIC (2024)
Our staff are regularly trained to recognise and prevent digital fraud risks such as phishing emails, synthetic media, and crypto scams.	Garanina, Klamm and Street (2022)
Our organisation has clear policies governing blockchain auditing, smart contract validation, and cryptocurrency transaction oversight.	Global Initiative (2022)
Our organisation invests in digital forensic capacity (e.g. dedicated IT forensic teams, outsourcing to forensic specialists).	Zhang, Ma and Meng (2025)
Our organisation complies with international fraud and audit standards that explicitly address digital risks (e.g. ISA 240).	IAASB (2021)

Appendix 2

Study-Level Coding Register

No	Author(s)	Year	Title (short)	Type	Findings	Gap	Human Capability (Binary)	Digital Capability (Binary)
1	Cressey, D.R.	1953	Other People's Money	Monograph	Identified pressure, opportunity, rationalisation	No digital/technological dimension included	1	0
2	Wolfe, D.T. & Hermanson, D.R.	2004	The Fraud Diamond	CPA Journal (Practitioner)	Lacked digital/AI context	Introduced capability construct for DFD	1	0
3	Dorminey, J. et al.	2012	The evolution of fraud theory	Issues in Accounting Education (SSCI)	No integration of digital fraud risks	Supports conceptual evolution for DFD	1	0
4	Trompeter, G. et al.	2013	Synthesis of fraud research	Auditing: A Journal of Practice & Theory (SSCI)	Did not examine digital fraud environment	Anchors empirical support for DFD constructs	1	0
5	Free, C.	2015	Looking through the fraud triangle	Meditari Accountancy Research (Scopus)	Ignored digital disinformation/fraud	Reinforces pressure/rationalisation debate for DFD	1	0
6	Schuchter, A. & Levi, M.	2016	The fraud triangle revisited	Security Journal (Scopus)	Neglects digital overlap issues	Highlights construct ambiguity addressed in DFD	1	0
7	Murphy, P.R. & Dacin, M.T.	2011	Psychological pathways to fraud	Journal of Business Ethics (SSCI)	Not extended to AI/tech fraud	Strengthens rationalisation construct relevance	1	0
8	Albrecht, W.S. et al.	2012	Fraud Examination	Cengage (Textbook)	Did not address blockchain/AI fraud	Grounds DFD in forensic practice	1	0

9	Morales, J. et al.	2014	The construction of the fraudster	Accounting, Organizations and Society (SSCI)	Did not apply to digital contexts	Strengthens rationalisation pressures in DFD	1	0
10	Free, C. & Murphy, P.R.	2015	The Decision to Co-Offend in Fraud	Contemporary Accounting Research (SSCI)	Neglects digital capability overlap	Clarifies opportunity-capability boundary for DFD	1	0
11	Jaakkola, E.	2020	Designing conceptual papers	Article	Outlined four approaches to conceptual papers and criteria for rigor and contribution	Does not address domain-specific content such as fraud or digital technologies	0	0
12	Whetten, D.A.	1989	What constitutes a theoretical contribution?	Article	Established criteria for theory development including why, how, what, who, where, and when	Prescriptive rather than domain-specific guidance	0	0
13	Garanina, T.; Klam, B.K.; Street, D.L.	2022	Forensic accounting in the digital era	Review	Identified growth areas such as data analytics, blockchain tracing, and AI tools in forensic practice	Limited empirical validation of tools in low-resource settings	0	1
14	Rezaee, Z.; Wang, J.	2018	Forensic accounting education and technology	Article	Argued for curriculum updates to include digital forensics and data analytics	Slow adoption in traditional programmes	1	0
15	Vidal, F.R. et al.	2024	Vulnerability detection for smart contracts SLR	SLR	Catalogue of vulnerability types and detection approaches, with performance metrics	Lack of forensic accounting perspective linking exploits to financial crime	0	1

16	Raju, B.; Vinayakumar, R.; Poornachandran, P.	2025	Automated smart contract vulnerability detection	Empirical	Proposed methods improved detection rates compared to baselines	Operational integration into audit workflows remains untested	0	1
17	Zhang, Y.; Ma, Z.; Meng, J.	2025	Auditing in the blockchain: a literature review	Review	Identified gaps in audit standards, traceability tools and skill shortages	Limited cross-country empirical studies, especially in emerging markets	0	1
18	Azimi, A.	2025	Digital Platform Durability	Doctoral dissertation	Catalogued design best practices that reduce common exploits	Translation of patterns into audit criteria is nascent	0	1
19	De Filippi, P.; Hassan, S.	2021	Smart contracts analysis	Conceptual	Showed mismatch between legal contract notions and code execution	Need for cross-disciplinary governance frameworks	1	0
20	Tapscott, D.; Tapscott, A.	2017	Blockchain Revolution	Book	Posited transparency and efficiency benefits, alongside governance challenges	Optimistic tone underplays fraud exploitation vectors	0	0
21	Pianoschi, A.	2025	Revising ISA 240 for digital world	Conceptual	Proposed amendments to ISA 240 to include digital evidence and analytics	Empirical testing of proposals is limited	1	0
22	Lokanan, M.E.	2015	Critique of the fraud triangle	Article	Highlighted methodological weaknesses and misapplications	Did not consider technological evolution in fraud	1	0

23	Maras, M.-H.; Alexandrou, A.	2019	Deepfakes and detection	Review/Technical	Outlined AI methods for synthesis and emerging detection algorithms	Limited discussion of economic/forensic impacts	0	1
24	Maras, M.-H.; Hurst, C.	2021	Synthetic media and corporate verification	Article	Demonstrated vulnerabilities in video and voice-based authentication	Empirical evidence on financial losses sparse	0	1
25	Journal of Accountancy	2024	What AI can do for auditors	Practitioner article	AI improves anomaly detection and scales analytics	Limited guidance on integrating AI into audit methodology	0	0
26	AICPA & CIMA	2025	Fraud and emerging technologies guidance	Guidance	Recommended controls, training, and technology adoption	Implementation challenges for smaller firms	1	0
27	Chainalysis	2024	Geography of Cryptocurrency 2024	Report	Identified hotspots of crypto-enabled crime and trends	Limited forensic accounting interpretation	1	0
28	Global Initiative	2022	Mirror Trading International case	Investigative report	Documented deceptive practices and investor losses	Policy recommendations need formal evaluation	1	0
29	Deloitte	2024	Deepfake banking and AI fraud risk	Report	Rising incidence, recommended detection and staff training	Unclear cost-benefit for small institutions	0	1
30	Regula	2024	Deepfake fraud survey	Report	High prevalence, increasing sophistication of attacks	Non-academic sampling limits generalisability	0	1
31	TransUnion Africa	2025	AI-powered fraud insights H1	Report	Increase in biometric and synthetic identity fraud	Limited academic corroboration	0	1

32	Sumsub	2023	KYC and identity fraud in Africa	Report	Biometric spoofing and synthetic IDs rising	Commercial dataset limitations	0	1
33	PwC	2019	Global economic crime survey	Practitioner report	Cyber-enabled fraud is fastest-growing category	Self-reported survey biases	1	0
34	KPMG	2021	Fraud and cyber survey	Practitioner report	Organisations lack integrated cyber-fraud response	Limited actionable frameworks for SMEs	1	0
35	ISACA	2023	Blockchain audit guidance	Guidance	Outlined assurance techniques and evidence sources	Practical adoption barriers in resource-poor firms	0	0
36	Reuters	2024	Investigative pieces on deepfakes	Journalism	High-impact cases highlighted	Journalism not peer-reviewed research	1	0
37	The Guardian	2024	Hong Kong deepfake transfer case	Journalism	Described modus operandi and victim impact	Limited technical forensics in reporting	1	0
38	Wall Street Journal	2025	CEO impersonation scams rise	Journalism	Significant increase in high-value impersonation schemes	Lacks academic methodological analysis	1	0
39	Trend Micro	2024	Deepfake threat reports	Industry report	Technical vulnerabilities and sectors at risk identified	Primarily technical, less forensic accounting linkage	0	1
40	CrowdStrike	2024	AI adversary report	Industry report	Adversaries increasingly automate deception at scale	Limited discussion on accounting implications	1	1
41	ISACA / IIA	2022	Auditing AI and blockchain guidance	Guidance	Provided checklists and assurance approaches	Needs empirical validation in diverse contexts	0	0

42	Nature HSS Communications	2024	AI and societal impacts	Peer reviewed	AI amplifies systemic risks and can be weaponised for fraud	High-level analysis, less domain-specific	0	0
43	SSRN working papers collection	2022	DeFi audits and assurance papers	Working papers	Identified gaps in auditability and assurance models for DeFi	Preliminary, requires peer review	0	0
44	Bibliometric SLRs on forensic accounting	2021	Forensic accounting bibliometrics	Review	Growth in digital themes and geographic spread noted	Call for more cross-disciplinary studies	0	0
45	Kranacher, M.; Riley, R.; Wells, J.	2011	Forensic Accounting and Fraud Examination	Book	Outlined investigative techniques and case handling	Limited digital forensic content in earlier editions	1	0
46	Healy, P.; Palepu, K.	2003	The fall of Enron	Case study	Exposed weak governance, opacity, and incentive structures	Pre-digital case, but instructive on cultural drivers	1	0
47	The DAO hack	2016	DAO smart contract exploit	Event case	Large funds drained due to code vulnerability	Lacked standard forensic approaches for smart contracts at the time	0	1
48	Mirror Trading International investigations	2020	MTI collapse reporting	Investigative reports	Revealed large-scale crypto Ponzi operations	Regulatory enforcement lagged	1	0
49	Qian, W.; Li, J.; Wang, T.; Chen, H.	2022	Detecting vulnerabilities in Ethereum smart contracts SLR	SLR	Surveyed tools and gaps in detection coverage	Interface with financial crime analysis missing	0	1

50	Azhar, S.; et al.	2023	AI-driven fraud detection in accounting	Article	AI models outperform traditional rules in pattern detection	Explainability and audit trail concerns remain	0	0
51	Smith, S.Z.	2024	How to audit smart contracts	Practitioner article	Outlined code review, simulation, and monitoring techniques	Scalability and standardisation challenges	0	0
52	Kumar, R.; et al.	2025	AI and disinformation in financial fraud	Article	AI enables rapid, targeted disinformation that can move markets	Empirical market impact estimation limited	1	0
53	Padalkar, M.	2023	NGO fraud risks and AI in Africa	Report	NGOs vulnerable to impersonation and payment scams	Limited sector-tailored mitigation strategies	1	0
54	OCCRP	2025	Identity fraud surge in Africa	Investigative report	Significant increases in synthetic identity cases tied to deepfakes	Quantitative estimation of economic loss incomplete	1	0
55	Heil, S.; Braun, P.	2022	Process mining for fraud detection	Article	Process mining uncovers deviations in workflows that signal fraud	Integration with forensic accounting practice is nascent	0	0
56	Kokina, J.; Davenport, T.H.	2017	Emergence of AI in accounting	Article	AI reshapes tasks, enabling predictive analytics and anomaly detection	Ethical and governance implications underexplored	1	0
57	Moll, J.; Yigitbasioglu, O.	2019	Analytics in audit	Article	Analytics improves risk assessment but requires skilled staff	Human-technology interplay needs more study	1	1

58	Leuz, C.; et al.	2003	Earnings management and fraud consequences	Article	Corporate incentives and weak governance correlate with fraud	Pre-digital focus, limited on tech-enabled manipulations	1	0
59	PwC	2022	Crypto and asset protection report	Practitioner report	Best practices for custody, monitoring, and controls	Regulatory fragmentation across jurisdictions	0	0
60	Kolk, A.; Levy, D.	2020	Blockchain governance and regulation	Article	Regulatory gaps create arbitrage opportunities that fraudsters exploit	Empirical studies on enforcement effectiveness lacking	1	0
61	Gao, L.; Zhao, Y.	2021	AI-based anomaly detection in finance	Article	High detection accuracy in controlled datasets	Real-world deployment and false positive handling challenges	0	0
62	Boehm, M.; et al.	2020	Smart contract formal verification	Article	Formal methods reduce but do not eliminate vulnerabilities	Complexity and scalability limit widespread adoption	0	0
63	Azar, G.; Dutta, S.	2021	Cybersecurity and financial fraud linkages	Article	Interconnected cyber- and fraud-risks create cascading failures	Few cross-disciplinary forensic studies	1	0
64	FSCA (South Africa)	2024	Warnings on deepfake investment scams	Regulatory report	Outlined incidents and recommended vigilance measures	Limited procedural guidance for forensic investigation	1	0
65	Global press coverage compendium	2019	High-profile fraud cases coverage	Journalism compendium	Showcases variety of fraud modalities and systemic impacts	Journalism lacks academic validation	1	0
66	FTX case analyses	2022	FTX collapse analyses	Case studies / analyses	Demonstrated weak governance, opacity and insider misuse	Novel crypto-specific controls under-researched	1	0

67	Wirecard scandal analyses	2020	Wirecard scandal analyses	Investigative articles	Complex schemes, regulatory blind spots, audit failures	Pre-digital aspects but relevant to organisational opportunity	1	0
68	Albrecht, W.S. et al.	2020	Advanced fraud detection techniques	Article	Outlined machine learning, network analysis, and process mining	Operational implementation barriers in practice	1	0
69	Zuboff, S.	2019	Surveillance capitalism	Book	Data-driven systems create new power imbalances and risks	Not directly focused on financial fraud	0	0
70	Broom, N.; Smith, L.	2021	Regulatory responses to crypto fraud	Article	Varied approaches, need for harmonisation	Implementation and enforcement effectiveness underexplored	1	0
71	Research working papers compendium	2023	Emerging AI fraud methods	Working papers	Novel attack vectors identified, requires rapid research response	Preliminary findings require validation	0	1
72	Bibliometric studies on forensic accounting	2022	Field mapping of forensic accounting	Review	Growth in digital topics and geographic diversification	Call for integration of technology-centred theory	0	0

Jovana Milenović
University of Niš, Faculty of
Economics
ORCID: 0000-0001-9718-0383

Ružica Petrović
University of Niš, Faculty of
Economics
ORCID: 0000-0002-7427-6486

LEGAL RISK AND FIRM MARKET VALUE: AN ANALYSIS OF THE RELATIONSHIP BETWEEN LITIGATION AND TOBIN'S Q RATIO OF BELEX15 INDEX COMPANIES

UDK 347.9:658.14

DOI <https://doi.org/10.54318/czf.2026.jm.024>

Abstract

In the contemporary business environment, measuring and quantifying legal risk, particularly litigation risk, is becoming increasingly important. As uncertainty about litigation outcomes and their potential impact on business operations grows, attention has shifted to how IAS 37 accounts for this risk and how capital markets assess firms' exposure to litigation. This paper aims to examine the accounting framework for recognizing and reporting legal risk under IAS 37 and to examine the relationship between expected and realized legal risk and the market valuation of firms, measured by Tobin's Q ratio, using a sample of nine entities included in the BELEX15 stock market index over the period from 2023 to 2025. The results of the analysis show that: 1) there is a strong and statistically significant negative correlation between expected litigation risk (ELR) and Tobin's Q ratio, indicating an inverse relationship between the amount of disclosed litigation provisions and firms' market value; 2) the correlation between realized litigation risk (RLR) and Tobin's Q indicates a moderate and statistically significant negative relationship at the 10% significance level, denoting a weaker, yet still present, association between firms' market value and the actual amounts of provisions utilized in connection with concluded litigation proceedings; and 3) there is a strong and statistically significant positive correlation between expected litigation risk (ELR) and realized litigation risk (RLR).

Key words: IAS37, litigation risks, Tobin's Q ratio, entity valuation, BELEX15

JEL Classification: M41, M42, K20, G32

Introduction

Legal risk, particularly that arising from litigation, is one of the most difficult types of business risk to measure. Unlike market or credit risk, which can be assessed using historical data and statistical models, litigation risk is characterized by uncertainty regarding the outcome, duration, and financial consequences of legal proceedings. Litigation may arise from contractual disputes, product liability, employment

disputes, intellectual property infringements, or environmental issues, and its outcome may remain uncertain for years, complicating both accounting recognition and market valuation.

The measurement of litigation risk is relevant to various stakeholders (Portela et al., 2025). Quantitative assessment can support investment and risk management decisions and provide legal advisers and damages experts with a basis for comparing subjective judgments about litigation outcomes with quantitative estimates. More accurate assessment may also contribute to reducing economically unjustified claims and litigation costs.

Litigation risk is closely related to the accounting framework established by International Accounting Standard 37 – Provisions, Contingent Liabilities and Contingent Assets (IAS 37). Liabilities arising from litigation are recognized as provisions when an outflow of resources is probable and can be reliably estimated, while uncertain obligations may be disclosed as contingent liabilities. Since IAS 37 requires considerable managerial judgment in assessing litigation outcomes and potential losses, an important question is whether financial statements adequately reflect firms' actual exposure to litigation risk and whether the market incorporates this exposure into firm valuation.

Empirical evidence indicates that litigation can significantly affect firm value. Billings et al. (2026) find that 23% of defendant firms experience a decline in market value exceeding 10% of assets following the filing of a lawsuit, while Sato et al. (2024) report that climate litigation is associated with an average firm value reduction of 0.41%. Litigation risk has also been shown to affect the cost of capital and stock liquidity, suggesting that markets may incorporate legal risk into valuation before litigation is resolved (Oliviero et al., 2024).

Tobin's Q captures the market valuation of a firm relative to the book value of its assets and reflects investors' expectations regarding future cash flows. It may therefore incorporate expectations concerning unresolved litigation, reputational consequences, and potential future obligations that are not fully reflected in accounting measures.

The aim of this paper is to examine whether a statistically significant relationship exists between litigation risk and firm value, measured by Tobin's Q, and determines the direction and strength of this relationship for companies included in the BELEX15 index. Litigation risk is represented by two indicators: expected litigation risk (ELR) and realized litigation risk (RLR), with the latter reflecting the financial consequences of realized litigation-related obligations.

To achieve this aim, the paper is organized as follows: it begins by examining IAS 37, with a special focus on its shortcomings in the area of litigation. Next, the paper examines the terminological and substantive differences between legal risk and litigation risk, since these factors are relevant determinants of firm value. The third

section covers the research methodology and empirical results. The final section summarizes the study's key findings.

1. Accounting treatment of litigation under IAS 37: a critical review

International Accounting Standard 37 – Provisions, Contingent Liabilities and Contingent Assets (IAS 37) was first issued by the International Accounting Standards Committee in 1998 and adopted by the International Accounting Standards Board (IASB) in 2001 as part of the International Financial Reporting Standards (IFRS) (IASB, 2001). The standard was introduced in response to accounting practices that had allowed companies to recognize provisions for future costs and obligations that did not represent present obligations. In the Republic of Serbia, IAS 37 is applied as part of the translated and officially published IFRS, including by companies whose securities are listed on an organised capital market, such as those included in the BELEX15 index.

IAS 37 establishes recognition and measurement requirements for provisions, contingent liabilities, and contingent assets. A provision is defined as a liability of uncertain timing or amount and is recognized when: (1) the entity has a present legal or constructive obligation resulting from a past event; (2) an outflow of resources embodying economic benefits is probable; and (3) the amount can be reliably estimated (IAS 37, para. 14). If these conditions are not met, the obligation may instead be disclosed as a contingent liability, depending on the probability of an outflow of resources (IAS 37).

Recognized provisions are measured at the best estimate of the expenditure required to settle the present obligation at the reporting date. For a large population of similar items, the expected value method is applied, while for a single obligation, such as an individual litigation proceeding, the most likely outcome generally represents the best estimate. When the time value of money is material, the provision is discounted to its present value. In litigation cases, however, these requirements involve considerable judgment by management and legal advisors regarding both the probability and financial consequences of an uncertain outcome.

Litigation is therefore a typical source of uncertainty addressed by IAS 37. When an unfavorable outcome is probable and the resulting loss can be reliably estimated, the obligation is recognized as a provision. When the outcome remains uncertain or cannot be reliably measured, it is disclosed as a contingent liability without recognition in the statement of financial position. This distinction between recognized provisions and contingent liabilities disclosed only in the notes is particularly relevant to this study, since litigation may affect firm value regardless of whether, and to what extent, it is reflected in the statement of financial position.

Despite improving the treatment of uncertain obligations, IAS 37 leaves considerable room for managerial judgment, particularly in assessing the probability of resource outflows and determining the best estimate of an obligation. The IASB has

acknowledged that recognition and measurement requirements may lead to different interpretations in practice and that more precise criteria could improve comparability (IASB, 2015). Haar et al. (2025) similarly argue that insufficiently precise requirements may reduce the reliability of provisions as indicators of future cash outflows. Using an option-based approach for companies listed on the London Stock Exchange, they identify widespread excessive provisioning and potential opportunities for earnings management through discretionary provision amounts.

The extent to which capital markets incorporate information on provisions and contingent liabilities has also been examined through the value relevance of accounting information. Lopes and Reis (2019), using companies from Portugal and the United Kingdom, find that recognized provisions have statistically significant value relevance for share prices, whereas contingent liabilities disclosed only in the notes generally do not, except for companies with an established risk management committee. This suggests that the distinction between recognition and disclosure may influence the extent to which investors incorporate legal risk into firm valuation and supports the use of Tobin's Q as a market-based measure of firm value complementary to IAS 37 accounting information.

The consistent application of IAS 37 may be particularly challenging in emerging markets characterized by weaker financial reporting oversight and less liquid capital markets. Acar and Ozkan (2017), examining 1,078 companies listed on the Istanbul Stock Exchange, find that 62% recognize provisions and only 32% fully comply with IAS 37 disclosure requirements. Compliance is positively associated with firm size, board independence, and ownership concentration. These findings suggest similar disclosure difficulties may arise in the BELEX15 market, as reflected in the inability to precisely determine the realized litigation risk (RLR) indicator for certain companies because of incomplete information in the notes to the financial statements.

Overall, IAS 37 provides an essential framework for accounting for legal risk but does not necessarily capture firms' full litigation exposure. Managerial discretion, differences between recognition and disclosure, and uneven disclosure practices indicate that accounting-based measures alone may not be sufficient for a comprehensive assessment of litigation risk.

2. Legal risk and litigation risk as determinants of firm value

The dynamic business environment requires the continuous evolution of legal rules governing companies' operations. As companies may deviate from expected requirements and conduct, legal risk must be analyzed and managed like other financial risks. Regulatory compliance risks, litigation risks, and other legal issues can significantly affect a company's operations and reputation (Guo & Tao, 2025).

The literature provides various definitions and manifestations of legal risk. It is commonly associated with an entity's risk of incurring liability for non-compliance

with applicable laws and regulations. More simply, legal risk is the intersection of a legal rule and an uncertain event affecting an entity's value (El Mir, 2022). For a legal rule to become a specific legal risk, an event must trigger liability or other legal consequences. Legal risk should not be defined as any risk involving legal professionals, since legal measures can also mitigate risks that are not inherently legal. Nor should it be equated exclusively with litigation risk, which most commonly concerns disputes between companies and other stakeholders (Ryu, 2024). Legal risk is broader and includes potential financial losses arising from changes in laws or regulations, non-compliance with external standards, difficulties in contract enforcement, and judicial and administrative proceedings (Kaup, 2024).

Despite its importance, legal risk is difficult to measure precisely and is generally not presented as a separate item in financial statements. Its manifestation depends on changes in legislation and regulations, regulatory authorities' decisions, and how companies apply regulatory requirements (Ryu, 2024). This study focuses specifically on litigation risk and its relationship with the value of companies included in the BELEX15 index. Litigation may adversely affect relationships with stakeholders, reputation, and business operations. Complex proceedings may last for years and generate substantial financial losses. Increased uncertainty may also cause investors and creditors to perceive the company as riskier, potentially limiting access to external financing and increasing borrowing costs. More restrictive financing conditions may constrain profitable investments, giving litigation both immediate financial consequences and longer-term effects on investment, revenues, and overall performance (Malm et al., 2023).

Litigation risk refers to uncertainty regarding whether a company will become involved in legal proceedings and consequently incur litigation costs, damages, or other financial obligations. Companies may face claims from investors, customers, employees, business partners, and regulatory authorities. Significant areas of litigation include patents and other intellectual property rights, antitrust violations, and shareholder class actions related to securities (Bennett et al., 2026). Environmental protection issues and liability for defective or harmful products have also emerged as important sources of litigation (Billings et al., 2026).

Litigation risk is therefore an important component of legal risk and of the broader management of non-financial risks. For publicly listed companies, litigation exposure may signal problems related to business operations, corporate governance, or regulatory compliance. Legal proceedings may increase uncertainty about future operations, prompting investors and creditors to reassess the company's creditworthiness, financial stability, and financing needs, even when its core business remains relatively stable (Ma et al., 2026).

3. The relationship between litigation risk and firm market value: an empirical analysis of Belex15 Index companies

To examine the relationship between expected and realized litigation risk and firm market value, the study used a sample of nine companies included in the BELEX15 stock market index as of the latest revision dated June 30, 2026. Table 1 presents the sample composition and basic information on each company.

Table 1: Characteristics of BELEX15 Companies

Company	Industry	Ticker Symbol
<i>Aerodrom Nikola Tesla a.d. Beograd</i>	Air Transport	AERO
<i>Dunav osiguranje a.d. Beograd</i>	Insurance	DNOS
<i>NIS a.d. Novi Sad</i>	Oil and Gas	NIIS
<i>Messer Tehnogas a.d. Beograd</i>	Industrial and Medical Gases	TGAS
<i>Metalac a.d. , Gornji Milanovac</i>	Manufacture of Cookware and Metal Products	MTLC
<i>Jedinstvo a.d. Sevojno</i>	Construction (Energy Infrastructure)	JESV
<i>Fintel energija a.d. Beograd</i>	Electricity Generation (Wind/Solar)	FINT
<i>Impol Seval a.d. Sevojno</i>	Aluminum Production	IMPL
<i>Energoprojekt holding a.d. Beograd</i>	Construction and Engineering	ENHL

Source: Author's own compilation based on available data.
<https://bgdx.rs/trgovanje/indeksi/belex15/korpa>

The study covers a three-year observation period from 2023 to 2025. The selection of this period was based on the following criteria: 1) availability of financial statements, as 2025 is the most recent fiscal year for which complete audited financial statements had been published at the time of the study; 2) a period of relatively normalized market conditions, as the observed years represent a period following the relative stabilization of market conditions affected by the COVID-19 pandemic and subsequent energy and inflationary shocks; 3) a sufficient number of observations, as the three years and nine companies generate up to 27 observations, thereby increasing the statistical precision of the correlation analysis compared with an analysis based on a single fiscal year; and 4) the temporal dynamics of litigation risk, given that legal proceedings generally extend over several years, making a multi-year analysis appropriate for capturing both the emergence of litigation and its eventual resolution.

The data were collected from financial statements available on the Serbian Business Registers Agency (APR) website, from which we obtained data on assets, liabilities, and equity, along with information disclosed on provisions and contingent liabilities in accordance with IAS 37. We obtained additional data from the Stock Analysis website.

Table 2: Indicators of Firm Value and Litigation Risk

Indicators	Measurement Method	Data Source
TobinQ	$\text{TobinQ} = \frac{\text{Market capitalization} + \text{Total liabilities}}{\text{Total assets}}$	Statement of Financial Position and Stock Analysis website
Expected Litigation Risk (ELR)	$\text{ELR} = \frac{\text{Provisions for litigation costs}}{\text{Total assets}}$	Statement of Financial Position and Notes to the Financial Statements
Realized Litigation Risk (RLR)	$\text{RLR} = \frac{\text{Utilized provisions for litigation costs}}{\text{Total assets}}$	Statement of Financial Position and Notes to the Financial Statements

Source: Author's research

Based on the theoretical foundations defined above, we formulate the following hypothesis (H1): Litigation risk indicators are statistically significantly related to Tobin's Q ratio. Before conducting the correlation analysis, we calculated descriptive statistics for all variables used in the study at the overall sample level. Table 3 presents the results.

Table 3: Descriptive Statistics of Variables for the Period 2023–2025

Indicators	N	Min	Max	Mean	Median	Std.	Skew	Kurtosis
TobinQ	27	0.0135	0.7832	0.3599	0.3097	0.26	0.30	-1.29
ELR	27	0.0004	0.0020	0.0011	0.0010	0.00	1.15	-0.07
RLR	27	0.0001	0.0001	0.0001	0.0001	0.00	2.02	2.36

Source: Author's research

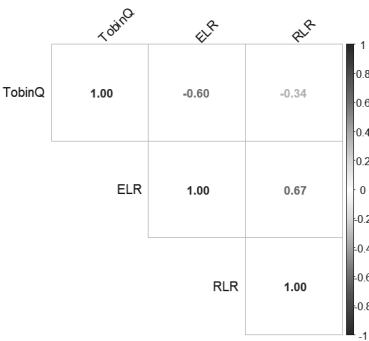
Tobin's Q ratio ranges from 0.0135 to 0.7832, with a mean of 0.36 and a median of 0.31. The slight positive skewness (skewness = 0.30) and negative kurtosis (−1.29) indicate a distribution that does not deviate substantially from normality. At the same time, the relatively high dispersion (standard deviation = 0.26) reflects significant differences in market valuation among the companies in the sample. ELR exhibits considerably lower and more narrowly distributed values (0.0004 to 0.0020), with moderate positive skewness (1.15), indicating that most companies have a low level of this indicator. In contrast, a small number of companies with more pronounced disclosed contingent liabilities shift the distribution toward higher values. RLR

exhibits the most pronounced deviation from a normal distribution, with a minimal difference between the minimum and maximum values, high positive skewness (2.02), and pronounced kurtosis (2.36). This indicates that most observations are concentrated around zero, with only a few exceptions showing higher values.

This distribution of the RLR indicator is partly attributable to limitations in data availability. For certain companies in the sample, it was not possible to determine the exact value of realized litigation risk because the notes to the financial statements did not disclose all the information needed for its calculation, specifically the amount of provisions utilized for concluded litigation proceedings during the observed year. We retained zero values for the RLR indicator because we operationalize realized litigation risk as the use of litigation-related provisions in a given year. Accordingly, when the available financial statements and accompanying notes identified no utilization of these provisions, we assigned a value of zero. This approach also allowed us to preserve the full sample of $N = 27$ and avoid disproportionately retaining only observations in which realized litigation risk was explicitly reported. Nevertheless, the absence of identified utilization does not necessarily imply the absence of actual litigation risk and may partly reflect differences in the transparency and detail of corporate disclosures. This methodological choice may also contribute to the low variability of the RLR indicator and is therefore acknowledged as a limitation when interpreting the RLR-related results.

ELR and, particularly, RLR exhibit pronounced deviations from a normal distribution. At the same time, Tobin's Q is relatively close to normality; Spearman's rank correlation coefficient was consistently applied to all variables. Given the relatively small sample of nine companies and a maximum of 27 firm-year observations, the analysis is limited to examining the direction and strength of the associations between litigation risk indicators and firm value. Spearman's rank correlation was selected as an appropriate non-parametric method that does not require the assumption of normality and is suitable for identifying monotonic relationships among the analyzed variables. Accordingly, the analysis is exploratory in nature and does not aim to establish causal relationships. The correlation matrix between the litigation risk indicators and Tobin's Q was calculated in R, following the methodological approach defined in the preceding subsections. The correlation coefficient r measures the strength of the relationship and ranges from -1 to $+1$ (Gujarati, 2004; Bhattacharjee, 2012). The closer the coefficient is to 1 in absolute terms, the stronger the relationship. Cohen (1988) classifies a correlation as small when r is between 0.10 and 0.29, medium when it is between 0.30 and 0.49, and large when it ranges from 0.50 to 1.00. Figure 1 presents the correlation results.

Figure 1: Correlation Matrix of the Selected Variables in the Study



Source: Author’s research

The correlation analysis, based on Spearman’s rank correlation coefficient because some variables deviated from normality, revealed a *strong, statistically significant negative correlation* between ELR and Tobin’s Q ($r = -0.60$; $p = 0.002$). This indicates that the market assigns lower valuations to firms with higher levels of disclosed expected litigation risk, suggesting that investors incorporate information on provisions and contingent liabilities into the firm valuation process even before the outcome of litigation is finally determined. A moderate, *statistically significant negative correlation* was identified between RLR and Tobin’s Q ($r = -0.34$; $p = 0.075$), indicating that the market responds less strongly to the realized effects of concluded litigation than to their expected, still uncertain component. This finding is partially consistent with the efficient market hypothesis, which holds that new risk information is incorporated into share prices as it emerges, rather than only when litigation is ultimately resolved. A *strong, statistically significant positive correlation* was found between ELR and RLR ($r = 0.67$; $p = 0.001$). This relationship can be explained by IAS 37, which governs the recognition and measurement of provisions, including provisions for litigation. Litigation provisions represent an accounting estimate of an expected future loss, based on an assessment of the probability that the risk will materialize and its potential financial impact.

Conclusion

This study examines the relationship between litigation risk and firm value, motivated by the discrepancy between the accounting treatment of litigation under IAS 37 and firms’ actual exposure to litigation risk. Using a sample of nine BELEX15 companies over the 2023–2025 period, the study applies correlation analysis to examine the relationship between expected litigation risk (ELR), realized litigation risk (RLR), and Tobin’s Q. The results support H1, indicating a negative relationship between litigation risk and firm value, although its strength differs between expected and realized litigation risk. The market appears to react more strongly to expected, still uncertain litigation risk than to its realized effects.

The findings have implications for several stakeholders. For investors and financial analysts, information on provisions and contingent liabilities represents a relevant source for assessing firm risk and value, despite the managerial discretion involved in their estimation. For insurance companies and damages experts, combining accounting-based litigation risk indicators with market-based measures may support investment and advisory decisions. For regulators and standard-setters, greater consistency in applying IAS 37 and more precise disclosure criteria could enhance the informational value of financial statements.

The study has several limitations. First, the sample of nine companies and 27 firm-year observations limits the generalizability of the findings. Second, RLR could not be precisely determined for certain companies due to insufficient disclosure in the notes to the financial statements. These observations were coded as zero, which may underestimate realized litigation risk and affect the results. Third, correlation analysis identifies associations between the variables but does not establish causality.

Future research should extend the sample by including a larger number of companies and a longer observation period, which would allow the panel structure of the data to be explicitly taken into account. The application of panel econometric techniques, such as fixed- or random-effects regression models with clustered standard errors, could provide more robust evidence on the relationship between litigation risk and firm value while controlling for unobserved firm-specific heterogeneity.

References

- Acar, E., Ozkan, S. (2017). Corporate governance and provisions under IAS 37. *EuroMed Journal of Business*, 12(1), 52-72. doi: 10.1108/emjb-03-2016-0007
- Bennett, B., Ham, C., Milbourn, T., Wang, Z. (2026). Firm-level litigation risk: Measurement and effects. *Management Science*. Advance online publication. doi: 10.1287/mnsc.2024.05821
- Beogradska berza. Available at: <https://bgdx.rs/trgovanje/indeksi/belex15/korpa>
- Billings, M. B., Holthausen, R. W., Petrovits, C., Wang, D. (2026). Incidence, risk, and disclosure of corporate litigation: Insights from federal court filings. *Journal of Accounting Research*. doi: 10.1111/1475-679x.70071
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- El Mir, N. (2022). *Deconstructing the construction of risk in organizations: the case of legal risk* (Doctoral dissertation, Université Paris Dauphine-PSL).
- Gujarati, D. N. (2004). *Basic Econometrics* (4th ed). Boston: The McGraw-Hill Companies
- Guo, H., Tao, L. (2025). The interplay between investor relations and corporate legal risks. *Finance Research Letters*, 108312. doi: 10.1016/j.frl.2025.108312

- Haar, L., Elharidy, A., Gregoriou, A. (2025). Marking-to-market IAS 37 provisions using options: an empirical demonstration. *Journal of Financial Reporting and Accounting*, 23(1), 387-403. doi: 10.1108/jfra-08-2022-0280
- IASB. (2015). Possible problems with IAS 37 (Agenda Paper 14B). International Accounting Standards Board. Available at: <chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.ifrs.org/content/dam/ifrs/meetings/2015/july/iasb/research-ias-37/ap14b-possible-problems-ias3.pdf>
- Kaup, S. (2024). Managing the legal risk of a business client. *European Research Studies Journal*, 27(3), 869-877. doi: 10.35808/ersj/3802
- Lopes, A. I., Reis, L. (2019). Are provisions and contingent liabilities priced by the market? An exploratory study in Portugal and the UK. *Meditari Accountancy Research*, 27(2), 228-257. doi: 10.1108/medar-09-2017-0212
- Ma, H., Ma, X., Lei, X., Jiang, Y., Li, Y. (2026). When lawsuits speak louder: Litigation risk and corporate financing constraints. *Economic Analysis and Policy*, 89, 578–590. doi: 10.1016/j.eap.2025.12.020
- Malm, J., Soyeh, K. W., Kanuri, S. (2023). Litigation risk and corporate performance. *Journal of Behavioral and Experimental Finance*, 37, 100725. doi: 10.1016/j.jbef.2022.100725
- International Standards Boars (IAS 37). Provisions, Contingent Liabilities and Contingent Assets. Available at: http://demo.paragraf.rs/demo/combined/Old/t/t2010_10/t10_0408.htm
- Oliviero, T., Park, M., & Zou, H. (2024). Liquidity effects of litigation risk: evidence from a legal shock. *The Journal of Law and Economics*, 67(1), 103-141. doi: 10.1086/727345
- Portela, J., Schwartz, E. S., & Garcia, J. A. (2025). *Litigation risk and the valuation of legal claims: A real option approach* (No. w33790). National Bureau of Economic Research. doi: 10.3386/w33790
- Ryu, D. (2024). The pricing and economic impact of legal risk [Working paper]. Saïd Business School, University of Oxford. doi:10.2139/ssrn.4967369
- Sato, M., Gostlow, G., Higham, C., Setzer, J., Venmans, F. (2024). Impacts of climate litigation on firm value. *Nature Sustainability*, 7(11), 1461-1468. doi: 10.1038/s41893-024-01455-y

Nikola Vidović
University of Banja Luka, Faculty of
Economics
ORCID: 0009-0006-3901-0590

MACROECONOMIC RESILIENCE AND INSTITUTIONAL VULNERABILITY: PUBLIC SECTOR TRANSPARENCY, FORENSIC AUDIT, AND THE MITIGATION OF SYSTEMIC FRAUD

UDK 330.34:351.72

DOI <https://doi.org/10.54318/czf.2026.nv.025>

Abstract

European economies did not absorb the shocks of the past fifteen years equally. The global financial crisis, the COVID-19 pandemic, and the 2022 energy and inflation shock left deeper marks on some countries than on others, and the Western Balkans came through worse than the European Union core. This paper asks how much of that difference comes from the quality of public institutions rather than from economic structure, and whether the control of public-sector corruption, the area forensic audit and fiscal transparency work in, cushions output when a shock hits. The analysis uses an unbalanced panel of 32 European economies, the EU-27 plus five Western Balkan states, over 2004 to 2023, with two-way fixed effects and Driscoll-Kraay standard errors, interacting the World Bank's Control of Corruption measure with crisis years. The interaction is positive in every specification that pools the crisis years: economies with stronger corruption control lose less output during crises. It is significant for the financial and pandemic shocks, where an extra point of Control of Corruption goes with roughly one percentage point less contraction, and a specification that keeps all three shocks shows the 2022 energy shock working in the opposite direction, which is why the full-sample estimate is weaker. No within-country link appears between corruption control and the fiscal balance, so the hypothesis of a fiscal-space channel is not supported in the panel, though a cross-country version of it stays open. The results make a case for treating forensic accounting and audit capacity as part of macroeconomic infrastructure, while noting that the governance proxy measures broad corruption control rather than forensic-audit capacity, and that perception-based data and a short panel limit what can be claimed.

Key words: macroeconomic resilience, fiscal space, forensic audit, control of corruption, fiscal transparency, institutional quality, panel data

JEL Classification: E62; H83; E02; O43; E32; H63

Introduction

The years since the global financial crisis have been a kind of natural experiment in resilience for European economies. The same broad shocks moved across a shared market and a common monetary setting, but they hit output very differently from one country to the next. The gap was clearest between the established EU member states and the candidate economies of the Western Balkans, where recoveries tended to be slower and fiscal room thinner. Much of the literature explains these differences through economic structure: trade openness, sectoral mix, the depth of capital markets. This paper looks at a different explanation, rooted in the quality of public institutions, and in particular in the control of the fraud, corruption, and fiscal leakage that forensic accounting exists to catch.

The logic is simple. Fraud and corruption drain the state's resources before a shock arrives. Diverted revenue, rigged procurement, and unrecorded liabilities all shrink the buffer a government can use when a downturn comes. If that holds, the institutions of public-sector control, audit bodies, transparent budgeting, and the forensic-accounting work behind them, are more than an after-the-fact accountability function: how strong they are before a crisis shapes how deep the contraction runs during it.

The idea is tested on an unbalanced panel of 32 European economies, observed each year from 2004 to 2023. Resilience is measured as the output response to shared crisis years. Institutional quality is captured mainly by the World Bank's Control of Corruption estimate, with Government Effectiveness as a second measure. The method is two-way fixed effects with Control of Corruption interacted with crisis years, and standard errors robust to serial correlation and cross-sectional dependence. One point should be clear from the outset: the governance indicator measures broad corruption control, not forensic-audit or supreme-audit capacity as such, so what the paper tests directly is the link between corruption control and resilience, with the connection to forensic accounting standing as the mechanism behind it rather than something measured here. The contribution is twofold: to bring the resilience and corruption-control literatures together on one European panel that spans both the EU core and the Western Balkan candidates, and to test the crisis payoff of corruption control directly, through an interaction with crisis years, rather than reading it from a standard growth regression.

Three results follow. The resilience interaction is positive in every specification that pools the crisis years: higher corruption control goes with smaller crisis-year output losses. For the financial and pandemic shocks the effect is significant, about one percentage point of output per unit of Control of Corruption, while a specification that separates 2022 shows that shock pulling in the other direction. No within-country relationship appears between corruption control and the recorded fiscal balance, which places the mechanism in the output channel and in cross-sectional differences in fiscal space.

1. Theoretical framework and literature

1.1 Resilience as an operational concept

Resilience has a conceptual side and an operational one. Briguglio, Cordina, Farrugia, and Vella (2009) separate a country's inherent vulnerability, its exposure to outside shocks from features like small size and openness, from its resilience, the policy-shaped capacity to withstand or recover from those shocks. Their framework puts good governance and sound macroeconomic management among the things that let an economy absorb a hit, which gives a route from institutional quality to macroeconomic outcomes. Resilience in this sense is not the absence of shocks but the ability to keep output near its path when they occur, which is what an interaction between institutional quality and crisis years is built to capture.

Sondermann (2018) provides the operational model. Working with a panel of advanced economies, he shows that countries with better-functioning structures and institutions take smaller output losses from a common shock, and that the difference is large. His approach defines resilience as the country-specific response to a shared disturbance and relates it to structural and institutional traits. That is the basis for the interaction setup used here, where the effect of a crisis on output is allowed to depend on the control of corruption. Ezcurra and Rios (2019) apply the same idea to Europe. They find that the quality of government mattered for regional resilience during the Great Recession, with an added role for governance in neighbouring regions. Their work is the closest precedent for the question asked here; the present paper differs in scope, moving from EU regions to a country panel that includes the Western Balkan candidates, and in isolating corruption control rather than a composite quality-of-government index.

1.2 Institutions, corruption, and growth

That institutions shape long-run performance is well established. Mauro (1995) gave early cross-country evidence that corruption lowers investment and, through it, growth. Tanzi and Davoodi (1997) sharpened the expenditure-side mechanism, showing that corruption raises public investment while cutting its productivity, tilting spending toward large, opaque capital projects that pay rents rather than returns. Rodrik, Subramanian, and Trebbi (2002) ranked institutional quality above trade integration and geography in explaining income differences across countries. Together this work shows both that corruption is costly and that its costs run largely through the public finances, the same channel this paper expects to condition resilience. The expenditure-side result matters most here: if corruption lowers the productivity of public spending in normal times, it leaves the state with less effective capacity to respond when a shock forces spending up.

1.3 Corruption, transparency, and fiscal space

The fiscal side is central here. The IMF's (2019) Fiscal Monitor examined the fiscal costs of corruption. It documented that corruption weakens the state's ability to collect revenue and to spend efficiently, and that stronger fiscal institutions, such as transparency, audit, and procurement controls, go with better outcomes. That is the clearest statement of the leakage mechanism and a direct precedent for relating a governance measure to fiscal performance. Alt and Lassen (2006) come at it from transparency. They build an index of fiscal transparency and show that more transparency goes with lower debt and deficits, independent of partisan and electoral factors, in a setup that treats transparency as itself partly endogenous.

Measuring fiscal space needs a construct, and Kose, Kurlat, Ohnsorge, and Sugawara (2022) supply one, a cross-country database covering debt sustainability, balance-sheet composition, and market access, and they show how these measures behave around bad events. Their construct lets fiscal space be treated as measurable rather than asserted, and frames the outcome side of the corruption-to-resilience channel.

1.4 Audit institutions and forensic accounting

If corruption drains fiscal space, the institutions that catch and deter it should help protect it. Two literatures speak to this directly. Blume and Voigt (2011) give the first cross-country assessment of whether the design of supreme audit institutions matters, relating the court, Westminster, and board models to fiscal and governance outcomes. Their work connects the audit function to macro-fiscal results and offers a way to code audit strength. Olken (2007) supplies the causal micro-evidence. In a randomised field experiment in Indonesia, raising the chance of a government audit sharply cut missing expenditures, and it beat community-based monitoring. That result is the anchor for reading audit intensity as a cause of reduced leakage rather than a correlate. The distance between a clean field experiment and a country panel like this one is real, and the paper does not pretend to close it. What Olken's result provides is a defensible claim that the association measured here has a plausible causal reading in at least one direction: better auditing cuts leakage, and less leakage is what preserves the buffer.

This is also the point to be explicit about the empirical proxy. The Control of Corruption indicator used below measures perceived control of corruption across a country's institutions. It is not a direct measure of forensic-audit capacity, the independence of a supreme audit institution, or the reach of fraud-detection systems. Those capacities help produce a good corruption-control score, which is why the indicator is a reasonable proxy for the theme of this paper, but the mapping is indirect. The results below are evidence on corruption control, with the forensic-accounting interpretation resting on the mechanism rather than a separate measurement.

1.5 Hypotheses

The literature points to two testable claims. One is about the fiscal channel in normal times. The other is about the payoff during crises.

H1. Higher control of corruption goes with stronger fiscal outcomes, lower public debt and smaller deficits, in normal periods.

H2. The output cost of a common shock is smaller where control of corruption is higher. That is, the interaction between control of corruption and crisis periods is positive.

2. Data and variables

The sample is an unbalanced annual panel of 32 European economies over 2004 to 2023: the 27 EU member states and five Western Balkan economies, namely Albania, Bosnia and Herzegovina, Montenegro, North Macedonia, and Serbia. Kosovo is left out because its coverage across the governance and national-accounts series is too thin to include, which leaves the Western Balkan group at five countries and 640 country-year observations in total. The 2004 start balances the need for a pre-crisis baseline against the availability of Western Balkan data.

The dependent variable for the resilience work is annual real GDP growth, from the World Bank's World Development Indicators (WDI). The two fiscal outcomes used to test H1, general government gross debt and net lending or borrowing, both as shares of GDP, come from the IMF's World Economic Outlook (WEO), which uses one consistent definition across the EU and the Western Balkans. A single source is used for the fiscal series rather than splicing European and national ones, so the definition stays consistent across the two blocs. The main institutional regressor is the Control of Corruption estimate from the World Bank's Worldwide Governance Indicators (WGI). Government Effectiveness, from the same source, serves as an alternative. Because this indicator captures broad corruption control rather than audit capacity as such, the tests speak directly to corruption control and only indirectly to forensic audit. Kaufmann, Kraay, and Mastruzzi (2010) describe and validate these indicators, and their documentation of the reported margins of error informs the caveats attached to inference below.

The controls follow standard growth-empirics practice: log real GDP per capita, trade openness measured as exports plus imports over GDP, consumer-price inflation, gross fixed capital formation as a share of GDP, and population growth, all from the WDI. A crisis indicator equals one in 2009, 2020, and 2022, the years of the financial, pandemic, and energy shocks. A restricted version drops 2022, and a separate specification keeps all three years but splits the interaction between 2009/2020 and 2022. The choice of crisis years is deliberately simple. These are the years of the three common shocks in the sample, and defining the indicator on the calendar rather than on country-specific recession dates keeps the interaction easy to interpret and hard to tune. Coverage of the assembled data runs above 99 percent,

with the only real gaps in a few early Western Balkan inflation figures, which the fixed-effects estimator drops. Table 1 gives summary statistics by bloc.

Table 1: Descriptive statistics by bloc, 2004–2023

Variable (mean)	EU-27	WB-5
Control of Corruption	0.995	−0.406
Real GDP growth, all years (%)	2.27	3.08
Real GDP growth, crisis years (%)	−1.96	−0.97
Gross debt (% GDP)	62.4	49.1
Net balance (% GDP)	−2.56	−2.65

Notes: EU-27 is the 27 EU member states; WB-5 is the five Western Balkan economies (Albania, Bosnia and Herzegovina, Montenegro, North Macedonia, Serbia). Crisis years are 2009, 2020, and 2022. Control of Corruption is the WGI estimate, higher is better. N = 640.

Source: Author's calculations based on World Bank (WDI, WGI) and IMF (WEO) data.

3. Empirical strategy

The resilience hypothesis is tested with two-way fixed effects on the country-year panel. For country i in year t , real GDP growth is regressed on Control of Corruption, its interaction with the crisis indicator, and the controls, with country and year fixed effects, as in equation (1):

$$growth_{it} = \beta_1 \cdot CoC_{it} + \beta_2 \cdot (CoC \times Crisis)_{it} + \gamma' X_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

The country effects μ absorb time-invariant differences: geography, legal origin, the long-run level of institutions. Identification then comes from within-country variation over time. The year effects λ absorb common shocks. Because the crisis indicator is a function of the year, it is collinear with the year effects and is not separately identified in the two-way model, so the crisis main effect is read from an auxiliary specification without year effects. The coefficient of interest is β_2 . A positive value means a given crisis goes with a smaller output loss where corruption control is higher, which is what H2 says. H1 is tested by swapping in gross debt or the net balance as the dependent variable and regressing it on Control of Corruption and the controls under the same structure.

Inference uses Driscoll-Kraay (1998) standard errors, which hold up under heteroskedasticity, serial correlation, and cross-sectional dependence, all plausible in a panel of connected European economies. Three diagnostics are reported: a Hausman test comparing fixed and random effects, a Wooldridge test for serial correlation, and Pesaran's test for cross-sectional dependence. The Hausman test is reported for information only; its standard form assumes the random-effects estimator is efficient under the null, which does not hold under the Driscoll-Kraay covariance used for inference, so it is read as a guide rather than a formal specification test. To handle the worry that a perception-based corruption measure

might move with the current business cycle, a version with Control of Corruption lagged one year is also estimated.

4. Results

4.1 Corruption control and output resilience

Table 2 reports the resilience models. The interaction between Control of Corruption and the crisis indicator is positive in every specification that pools the crisis years, in line with H2: economies with stronger corruption control lose less output when a common shock hits. In the baseline two-way model (column B) the interaction is 0.55, though not individually significant once year effects absorb the common part of the crisis. The auxiliary model that shows the crisis main effect (column A) puts the average crisis-year contraction at about six percentage points, which gives the interaction's size some context.

The result that matters turns on 2022. Restrict crises to the financial and pandemic shocks of 2009 and 2020 (column E) and the interaction rises to 1.03, significant at five percent. An extra unit of Control of Corruption goes with about one percentage point less output contraction. A unit is a sizeable move: a little under the 1.2-point gap between the median EU and median Western Balkan economy, and somewhat larger than one pooled standard deviation of the indicator (0.86). The reading should stay cautious, since the index is perception-based and has no natural cardinal scale. The lagged-corruption (column C) and Government Effectiveness (column D) versions give interactions of nearly the same size, so the result is not an artifact of contemporaneous perception and not tied to one governance measure.

Two features of Table 2 are worth a note. The reported R^2 is that of the fixed-effects regression once the country and year effects are partialled out; it is modest, around 0.10 in the two-way models, because those effects already absorb most of the variation in growth, so it should not be read as the explanatory power of the interaction. The sample reconciles simply: 640 observations in principle, 636 once four early inflation values for Bosnia and Montenegro are dropped, and 606 in column C, where the one-year lag removes 2004 for all 32 countries and two of the four inflation gaps fall in 2005 ($640 - 32 - 2 = 606$). Every column covers all 32 countries.

Table 2: Control of corruption and output resilience (dependent variable: real GDP growth)

	A	B	C	D	E
Control of Corruption	0.835	−0.411			−0.468
	(1.074)	(1.258)			(1.278)
CoC × Crisis	0.554	0.552			
	(0.340)	(0.445)			
Crisis (main effect)	−6.034***	abs.	abs.	abs.	abs.
	(1.533)				
CoC (t−1)			−0.425		
			(1.151)		
CoC(t−1) × Crisis			0.516		
			(0.457)		
Gov. Effectiveness				−0.256	
				(0.767)	
GE × Crisis				0.520	
				(0.461)	
CoC × Crisis (2009/2020)					1.027**
					(0.452)
Controls / Country FE	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes
Observations	636	636	606	636	636
Countries	32	32	32	32	32
R ²	0.362	0.100	0.108	0.098	0.107

Notes: Driscoll-Kraay standard errors in parentheses. "abs." means absorbed by year fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The interaction term is the test of H2. R² is that of the fixed-effects regression after the reported effects are removed; column A includes entity effects only.

Source: Author's calculations based on World Bank (WDI, WGI) and IMF (WEO) data.

To show the 2022 effect directly rather than by dropping the year, Table 3 keeps all three shocks and interacts Control of Corruption separately with the 2009/2020 years and with 2022. The 2009/2020 interaction is positive and significant (0.998), while the 2022 interaction is negative and significant (−0.504). Corruption control cushioned the financial and pandemic shocks but, if anything, added to the loss in the energy shock. One reading is the greater exposure of high-governance economies to imported gas, but this is a conjecture: 2022 growth was still positive in most

economies, so the term may instead reflect uneven post-pandemic recoveries. Either way, the split removes any suggestion that the headline rests on a convenient choice of crisis years.

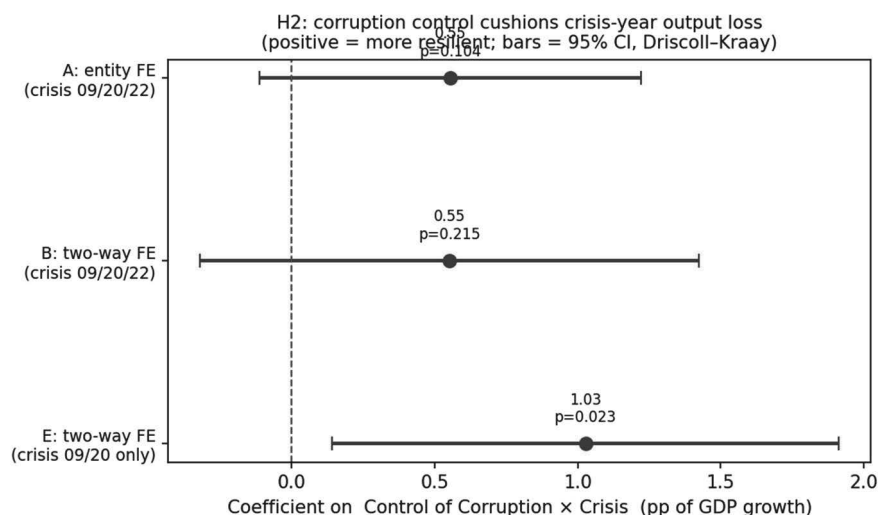
Table 3: Resilience with the 2022 energy shock separated (dependent variable: real GDP growth)

Term	Coefficient (SE)
Control of Corruption	−0.437 (1.276)
CoC × Crisis (2009/2020)	0.998 (0.448)**
CoC × Crisis (2022)	−0.504 (0.229)**
Controls / Country FE / Year FE	Yes
Observations	636
Countries	32
R ²	0.108

Notes: two-way fixed effects, Driscoll-Kraay standard errors in parentheses. ** $p < 0.05$.

Source: Author's calculations based on World Bank (WDI, WGI) and IMF (WEO) data.

Graph 1: The Control of Corruption × Crisis coefficient across specifications



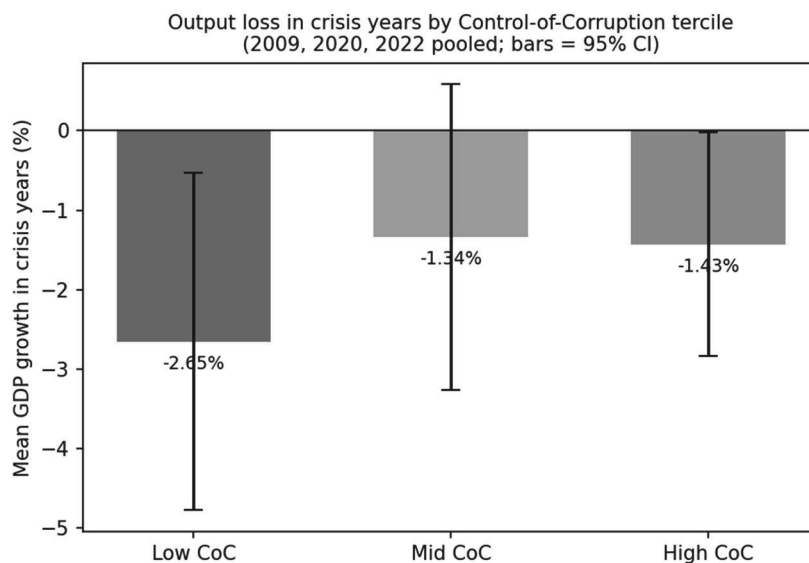
Note: 95% confidence intervals, Driscoll-Kraay. Each pooled-crisis estimate shown is positive; the 2009/2020 specification excludes zero.

Source: Author's calculations based on World Bank (WDI, WGI) and IMF (WEO) data.

Graph 1 shows the pattern, plotting the interaction across the main specifications with confidence intervals. Graph 2 shows the mechanism directly. Pooling the crisis

years, mean output loss falls from -2.65 percent in the lowest tercile of corruption control to -1.43 percent in the highest.

Graph 2: Mean crisis-year GDP growth by tercile of Control of Corruption



Note: crisis years 2009, 2020, and 2022 pooled; 95% confidence intervals.

Source: Author's calculations based on World Bank (WDI, WGI) and IMF (WEO) data.

4.2 The fiscal-space channel

Table 4 tests H1, and does not support it. Inside the fixed-effects structure, Control of Corruption is not significantly related to either gross debt or the net balance, whether entered contemporaneously or lagged, and the debt coefficient even carries the wrong sign. H1 should therefore be read as not supported within this specification, not as confirmed. The reason is not that the leakage mechanism is absent but that country fixed effects remove exactly the cross-country variation where the corruption-fiscal-space link is expected to sit. Over two decades within one country, Control of Corruption barely moves, while debt is driven by the crises themselves. Whether a cross-country relationship between institutional quality and fiscal space holds is left open here. In this sample the resilience effect works through output and through cross-country differences in fiscal space, not through year-to-year within-country fiscal adjustment. That fits the descriptives: the two blocs differ a lot in the level of corruption control (Table 1) even though their within-period fiscal paths do not track it.

Table 4: Control of corruption and fiscal space (two-way fixed effects)

Specification	Gross debt (% GDP)	Net balance (% GDP)
Control of Corruption	2.928 (2.502)	-0.454 (0.764)
Control of Corruption (t-1)	2.257 (2.102)	-0.746 (0.722)
p-value (contemporaneous)	0.242	0.552

Notes: Driscoll-Kraay standard errors in parentheses. Controls and country and year fixed effects included. N = 636 for the contemporaneous rows and 606 for the lagged rows. No coefficient is significant at conventional levels.

Source: Author's calculations based on World Bank (WDI, WGI) and IMF (WEO) data.

4.3 Heterogeneity and diagnostics

A specification that lets the resilience interaction differ by bloc estimated with the full set of lower-order terms, points to a buffer centred on the EU sub-panel, with an EU-bloc interaction of about 0.93. The Western Balkan interaction terms are individually insignificant and, given the five-country cross-section, too imprecisely estimated to support a numerical claim; the full specification and its standard errors are in the Appendix (Table A1). The diagnostics support the estimation choices: the Hausman test favours fixed effects over random effects ($\chi^2(7) = 30.4$), read with the caveat in Section 3. The Wooldridge test rejects no first-order serial correlation ($p < 0.001$), which is why robust inference is used. Pesaran's test does not reject cross-sectional independence once year effects are in ($CD = -1.31$, $p = 0.19$), so the year dummies already absorb the common shocks and Driscoll-Kraay errors remain appropriate on serial-correlation grounds.

5. Robustness

The main result holds up under several changes. Replacing contemporaneous Control of Corruption with its one-year lag leaves the interaction about the same (Table 2, column C), easing the worry that the measure just tracks the current state of the economy. Government Effectiveness gives a nearly identical interaction (column D), so the finding reflects a broad quality-of-government idea rather than one indicator. The contrast between crisis definitions is handled directly in Table 3 rather than by dropping a year, which answers the concern that the headline depends on the choice of crisis years. And since cross-sectional dependence does not appear once year effects are included, the results do not hinge on the choice between Driscoll-Kraay and entity-clustered errors.

6. Discussion and policy implications

The evidence supports a limited but useful claim. The control of public-sector corruption goes with greater macroeconomic resilience, and in the financial and

pandemic episodes the link is strong enough to matter both economically and statistically. Set next to Olken's (2007) micro-evidence and the IMF's (2019) work on fiscal costs, the most sensible reading is that the institutions of detection and deterrence, audit bodies, transparent budgeting, and the forensic-accounting work behind them, protect the fiscal and allocative integrity that lets an economy absorb a shock. That shifts forensic accounting from an after-the-fact accountability function to part of the machinery that carries a country through stress.

For forensic accounting specifically, the value is not confined to detecting and prosecuting fraud after the fact: the same capacity, applied continuously through public audit and transparent reporting, keeps the public balance sheet closer to its stated position, which is what leaves room to act when a shock arrives.

Two points temper the policy reading. First, in this sample the resilience effect runs through output, not through the recorded within-country fiscal balance, so audit and transparency reforms should be justified by what they do for the efficiency and credibility of the public sector, not by an expectation of lower measured deficits in the short run. Second, the effect reverses in 2022, a reminder that governance buffers are shock-specific: they cushion crises that spread through fiscal and financial channels better than an energy shock, though the 2022 sign may also reflect uneven post-pandemic recoveries rather than exposure alone. For the Western Balkans, where the accession path already stresses rule-of-law and public-financial-management reform, the results argue for building forensic-audit and transparency capacity early. They also caution that the returns are conditional, and below some level of institutional development they may be small.

Conclusion

Using a panel of 32 European economies over 2004 to 2023, this paper shows that stronger control of corruption goes with smaller crisis-year output losses, significantly so for the financial and pandemic shocks, and places the mechanism in the output channel rather than in within-country fiscal adjustment. The hypothesis of a within-country fiscal-space channel (H1) is not supported, while its cross-country version is left open. The results support treating forensic accounting and public-audit capacity as part of a country's resilience setup, with the proviso that the governance proxy stands in for those capacities rather than measuring them. Some limits apply. The Worldwide Governance Indicators are perception-based and measured with real error. The lagged specification eases but does not remove the resulting endogeneity, and the paper does not claim to have identified a structural causal parameter. The Western Balkan cross-section is small, which limits what can be said about bloc-level heterogeneity. The fixed-effects design, while conservative and appropriate, is by construction poor at picking up the slow-moving, cross-sectional part of the corruption-fiscal-space relationship. Several extensions follow. A between-effects or random-effects treatment of H1 would recover the cross-sectional fiscal channel. A local-projection version of H2 would trace state-dependent impulse responses in

the spirit of Sondermann (2018). A direct measure of forensic-audit or supreme-audit-institution strength, along the lines of Blume and Voigt (2011), would move the analysis from a governance proxy to the specific institutions this literature is about. A dynamic specification estimated by system GMM would be a natural robustness check on persistence, though it is not required for the result reported here.

References

- Alt, J. E., & Lassen, D. D. (2006). Fiscal transparency, political parties, and debt in OECD countries. *European Economic Review*, 50(6), 1403–1439. <https://doi.org/10.1016/j.eurocorev.2005.04.001>
- Blume, L., & Voigt, S. (2011). Does organizational design of supreme audit institutions matter? A cross-country assessment. *European Journal of Political Economy*, 27(2), 215–229. <https://doi.org/10.1016/j.ejpoleco.2010.07.001>
- Briguglio, L., Cordina, G., Farrugia, N., & Vella, S. (2009). Economic vulnerability and resilience: Concepts and measurements. *Oxford Development Studies*, 37(3), 229–247. <https://doi.org/10.1080/13600810903089893>
- Driscoll, J. C., & Kraay, A. C. (1998). Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics*, 80(4), 549–560. <https://doi.org/10.1162/003465398557825>
- Ezcurra, R., & Rios, V. (2019). Quality of government and regional resilience in the European Union: Evidence from the Great Recession. *Papers in Regional Science*, 98(3), 1267–1290. <https://doi.org/10.1111/pirs.12417>
- International Monetary Fund. (2019). *Fiscal Monitor, April 2019: Curbing corruption* (Chapter 2). International Monetary Fund. <https://www.imf.org/en/Publications/FM/Issues/2019/03/18/fiscal-monitor-april-2019>
- Kaufmann, D., Kraay, A., & Mastruzzi, M. (2010). *The worldwide governance indicators: Methodology and analytical issues* (Policy Research Working Paper No. 5430). World Bank. <https://ssrn.com/abstract=1682130>
- Kose, M. A., Kurlat, S., Ohnsorge, F., & Sugawara, N. (2022). A cross-country database of fiscal space. *Journal of International Money and Finance*, 128, 102682. <https://doi.org/10.1016/j.jimonfin.2022.102682>
- Mauro, P. (1995). Corruption and growth. *The Quarterly Journal of Economics*, 110(3), 681–712. <https://doi.org/10.2307/2946696>
- Olken, B. A. (2007). Monitoring corruption: Evidence from a field experiment in Indonesia. *Journal of Political Economy*, 115(2), 200–249. <https://doi.org/10.1086/517935>

Rodrik, D., Subramanian, A., & Trebbi, F. (2002). *Institutions rule: The primacy of institutions over integration and geography in economic development* (IMF Working Paper No. 02/189). International Monetary Fund. <https://www.imf.org/external/pubs/ft/wp/2002/wp02189.pdf>

Sondermann, D. (2018). Towards more resilient economies: The role of well-functioning economic structures. *Journal of Policy Modeling*, 40(1), 97–117. <https://doi.org/10.1016/j.jpolmod.2018.01.002>

Tanzi, V., & Davoodi, H. (1997). *Corruption, public investment, and growth* (IMF Working Paper No. 97/139). International Monetary Fund. <https://www.imf.org/external/pubs/ft/wp/wp97139.pdf>

Appendix

Table A1: Heterogeneity of the resilience effect by bloc (dependent variable: real GDP growth)

Term	Coefficient (SE)
Control of Corruption	−0.591 (1.335)
CoC × Crisis	0.931 (0.659)
CoC × WB-5	2.607 (2.718)
Crisis × WB-5	−2.236 (2.544)
CoC × Crisis × WB-5	−8.750 (6.196)
ln GDP per capita	2.202 (1.429)
Trade openness	0.034 (0.011)***
Inflation (CPI)	−0.036 (0.092)
Gross fixed capital formation	0.202 (0.091)**
Population growth	−0.316 (0.183)*
Observations	636
Countries	32
R ²	0.112

Notes: dependent variable is real GDP growth; two-way fixed effects estimated with the full set of lower-order terms; Driscoll-Kraay standard errors in parentheses. The WB-5 and Crisis main effects are absorbed by the country and year effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The Hausman, Wooldridge, and Pesaran diagnostics reported in Sections 3 and 4.3 refer to the baseline specification (Table 2, column B).

Source: Author's calculations based on World Bank (WDI, WGI) and IMF (WEO) data.

Željko Vujinović
Elektroprivreda Srbije (EPS),
Renewable Energy Sources Branch

A METHODOLOGICAL FRAMEWORK FOR INTERPRETING LEGAL PROVISIONS THAT REFER TO THE ACCOUNTING REGULATORY FRAMEWORK: A FORENSIC ACCOUNTING APPROACH

UDK 347.73:657.6

DOI <https://doi.org/10.54318/czf.2026.zv.026>

Abstract

Legal provisions that directly or indirectly refer to the accounting regulatory framework may raise a specific interpretative issue when the relevant accounting rules require professional accounting judgment. The aim of this paper is to develop an analytical framework that systematically links the legal norm, the accounting regulatory framework, professional accounting judgment, and the significance of its result for the application of law. The research is based on a normative analysis of Serbian legislation, an analysis of the relevant accounting regulatory framework, and a qualitative analysis of selected Serbian court decisions. The proposed framework is applied in detail to Article 7 of the Serbian Property Tax Act and Administrative Court judgment 9 U 10493/18, while its transferability is additionally examined through other judicial decisions. The forensic accounting approach examines the regulatory basis, accounting policies, available data, methods and assumptions, and supporting documentation in order to reconstruct the process by which the professional conclusion was reached. The results indicate that the proposed framework enables a methodologically clear connection between professional accounting judgment and the application of a legal norm.

Keywords: professional accounting judgment; forensic accounting; IAS/IFRS; legal provisions; accounting regulatory framework; fair value.

JEL Classification: M41; K34

Introduction

Modern financial legislation increasingly links legal consequences to accounting categories, financial information, and rules derived from the accounting regulatory framework. This occurs both when legislation expressly refers to International Accounting Standards (IAS), International Financial Reporting Standards (IFRS), or accounting policies and when it uses accounting categories that cannot be determined

without that framework. In such cases, an accounting result may become directly relevant to legal rights or obligations.

The issue becomes more complex when a principle-based accounting rule requires professional accounting judgment. Such judgment is not a mechanical application of a predetermined rule but an application of professional expertise to the facts and circumstances of the case (Ajekwe, 2022; Bennett et al., 2006). Where a legal provision assigns significance to its result, legal analysis may need to identify not only the accounting result but also the rules, facts, assumptions, and assessments on which it was based.

Professional accounting judgment and judgment and decision-making constitute a well-established stream of international accounting research. Existing studies examine professional judgment in financial reporting, auditing, and the application of accounting standards, including the role of knowledge, task characteristics, environmental factors, and the structure of accounting standards in professional judgment and decision-making (Bonner, 1999; Trotman, 1998; Bennett et al., 2006; Trotman et al., 2011; Mala & Chand, 2015; Ajekwe, 2022; Fomina et al., 2022; Krpić, 2023). While this literature provides an extensive basis for understanding professional accounting judgment, the literature reviewed for this study did not identify a methodological framework specifically designed to analyse situations in which a legal norm assigns significance to an accounting result that depends on such judgment. The present study addresses this intersection by systematically linking the legal norm, the accounting regulatory framework, professional accounting judgment, and the legal significance of its result.

Accordingly, the research asks: How can legal provisions that directly or indirectly refer to the accounting regulatory framework be analysed when professional accounting judgment affects their application? The study combines normative analysis of Serbian legislation and the accounting framework with qualitative analysis of selected court decisions. The framework is examined through Article 7 of the Property Tax Act and Administrative Court judgment 9 U 10493/18, with additional decisions used to test transferability.

The forensic accounting perspective reconstructs how conclusions based on professional judgment were reached. The scientific contribution is an analytical framework linking the legal norm, accounting regulation, and professional accounting judgment, with practical relevance for courts, authorities, lawyers, accountants, auditors, and forensic accountants.

1. Normative reference to the accounting regulatory framework

Legal provisions may connect with the accounting regulatory framework in different ways. A statute may expressly refer to IAS, IFRS, or an accounting framework established by accounting legislation, or attach a legal consequence to an accounting category determined under accounting rules. These forms are described as direct and

indirect normative reference. The distinction identifies how accounting enters the application of a legal norm, especially where policies, valuation methods, and professional judgment mediate between statutory wording and the legally relevant result.

1.1. Direct normative reference

Direct normative reference exists where a legal provision expressly requires the application of IAS, IFRS, or the accounting regulatory framework as a rule relevant to its application. The Serbian Accounting Act is the basic example because it requires specified categories of entities to apply IAS/IFRS in preparing financial statements. Regulations governing banks and the capital market also rely on IAS/IFRS directly or through the accounting framework established by accounting legislation, making accounting regulation part of the broader system of financial regulation.

1.2. Indirect normative reference

Indirect normative reference exists where a legal provision links a legal consequence to an accounting category or result determined under the accounting regulatory framework, such as financial result, capital, net assets, or asset value. The framework is then the professional basis for determining the category to which the law assigns legal significance.

Article 7 of the Serbian Property Tax Act is central to this research. For certain taxpayers, the tax base is the fair value of real estate recognised in the accounting records, provided that the property is measured using the fair value method in accordance with IAS/IFRS and adopted accounting policies (Property Tax Act, 2001, Art. 7). The statute expressly refers to IAS/IFRS while assigning legal significance to a fair value determined through that framework. The model classifies Article 7 as an indirect reference with an explicit IAS/IFRS link because an accounting category mediates between the framework and legal consequence. This classification concerns how the legally relevant result is produced, not whether the statutory text contains an express reference.

1.3. Normative reference in Serbian legislation

The connection between law and accounting is not limited to tax law. Table 1 presents characteristic examples from Serbian legislation and shows that normative reference may concern both the accounting framework itself and accounting categories produced through its application.

Table 1: Examples of normative reference to the accounting regulatory framework in Serbian legislation

Legislation	Form of reference	Relevant category/framework
Accounting Act (2019)	Direct	IAS/IFRS
Banking Act (2005)	Direct	IAS/IFRS
Capital Market Act (2021)	Direct	IAS/IFRS
Corporate Income Tax Act (2001)	Indirect	Financial result
Property Tax Act, Article 7 (2001)	Indirect*	Fair value
Companies Act (2011)	Indirect	Capital, net assets and financial statements
Bankruptcy Act (2009)	Indirect	Asset valuation and financial position

**Article 7 contains an explicit link to IAS/IFRS and accounting policies. Its classification as an indirect reference reflects the fact that the legal consequence depends on fair value recognised in the accounting records, whose content and method of determination arise from the relevant accounting regulatory framework.*

Source: Author, based on an analysis of relevant Serbian legislation.

The presence of an accounting category in a legal provision does not automatically mean that professional accounting judgment is required. Some accounting rules provide a sufficiently determinate answer; others require an assessment of facts and circumstances, a choice among permitted alternatives, or another professional conclusion. This difference leads to the next analytical step.

1.4. Relationship between normative reference and professional judgment

Once a legal provision is found to depend on the accounting regulatory framework, the relevant accounting rule must be identified. The central question is whether applying that rule requires professional accounting judgment. If not, no separate analysis of judgment is needed. If so, the manner in which judgment was exercised and the legal significance of its result must be examined. This distinction is the starting point of the framework below.

2. Analytical framework for identifying the role of professional judgment in the application of law

2.1. Professional accounting judgment

A legal provision may refer to accounting regulation without requiring professional judgment in every case. Judgment becomes necessary where a general accounting rule must be applied to specific facts through professional assessment, choice among permitted alternatives, or assumptions (Ajekwe, 2022; Fomina et al., 2022; Krpić, 2023). For this research, professional accounting judgment is the application of professional knowledge, experience, and the relevant accounting framework to reach a reasoned conclusion where proper application requires assessment, choice, or interpretation.

Professional judgment is not arbitrary; it is exercised within accounting rules and on the basis of information available when the conclusion is reached. It may be required when estimating useful life, determining impairment, selecting a valuation method, or assessing whether fair value changes require a new valuation (Ajekwe, 2022; IFRS Foundation, 2021a, 2021b, 2021c, 2021d). The process of judgment must also be distinguished from its result - a value, estimate, classification, accounting treatment, or other conclusion to which law may assign significance.

2.2. Forensic accounting approach

Where a legally relevant accounting result depends on professional judgment, it may be insufficient to identify only the final amount or accounting treatment. A forensic accounting analysis examines the regulatory basis, accounting policies, facts and data available at the relevant time, methods and assumptions used, and supporting documentation from which the decision process can be reconstructed (Honigsberg, 2020; Tekavčič & Damijan, 2021).

The forensic component reconstructs the decision trail and tests consistency among the rule, available facts, accounting policy, methods, assumptions, documentation, and reported result. It does not replace the original judgment retrospectively, but determines the basis of the existing conclusion and whether the process accords with the relevant accounting framework. Where law assigns legal effect to an accounting result, how that result was reached may therefore be legally relevant.

Professional accounting judgment may also result in different conclusions being reached by qualified professionals in relation to the same accounting or factual circumstances, particularly where the applicable accounting framework permits judgment and the available evidence reasonably supports more than one assessment. The existence of competing professional conclusions does not, by itself, establish that one of them is incorrect. In such circumstances, the forensic accounting approach may provide a structured basis for comparing the regulatory basis of each conclusion, the accounting policies applied, the relevance and reliability of the data used, the methods and significant assumptions adopted, the consistency of their

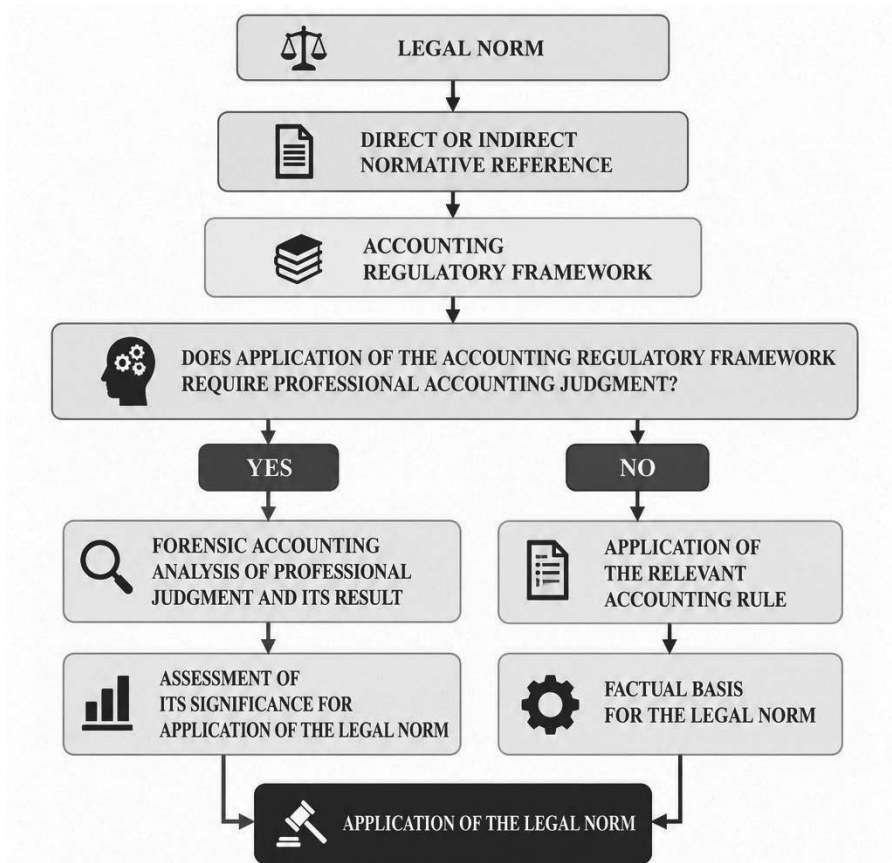
application, and the supporting documentation. Where competing professional assessments are presented in judicial proceedings, these elements may also assist the court in evaluating the reasoning and evidential support underlying each assessment, while the final evaluation of the evidence and the legal conclusion remain matters for the court.

The proposed forensic accounting reconstruction also has certain methodological similarities with procedures used in auditing accounting estimates. Research on the auditing of complex accounting estimates and fair value measurements similarly emphasizes the importance of critically evaluating the evidence, methods, assumptions and estimation uncertainty involved in such assessments (Griffith et al., 2015; Cannon & Bedard, 2017). In particular, ISA 540 requires auditors to evaluate the methods, significant assumptions and data used in making accounting estimates and to assess whether the resulting estimates and related disclosures are reasonable in the context of the applicable financial reporting framework. Depending on the audit approach, the auditor may also develop a point estimate or range against which management's estimate is evaluated (IAASB, 2018).

This methodological parallel does not mean that forensic accounting reconstruction is equivalent to an audit. The objectives and contexts are different. Audit procedures concerning accounting estimates are performed within the audit of financial statements and are directed toward obtaining sufficient appropriate audit evidence to support the auditor's conclusions. The forensic accounting approach proposed in this paper is instead directed toward reconstructing and documenting the professional reasoning underlying a legally relevant accounting result and examining its regulatory, methodological and evidential basis for subsequent legal analysis. The similarity therefore concerns certain analytical procedures, rather than the purpose or institutional function of the two approaches.

2.3. Proposed analytical framework

The proposed framework systematically connects the legal norm, the accounting regulatory framework, professional accounting judgment, the accounting result, and the legal significance of that result. Figure 1 presents the analytical sequence. Its key decision point asks whether application of the relevant accounting framework requires professional accounting judgment.

Figure 1: Analytical Framework

Source: Author.

If the answer is NO, the relevant accounting rule is applied and its result is used in applying the legal norm. If the answer is YES, the professional judgment and its result are analysed through the forensic accounting approach before their significance for the legal norm is assessed. This YES/NO decision point prevents professional judgment from being presumed whenever law relies on accounting, while still identifying situations in which the legally relevant accounting result was not produced solely by mechanical application of a rule.

2.4. Practical application of the analytical framework

For practical purposes, the framework can be expressed through the six questions shown in Table 2.

Table 2: Six questions for practical application of the analytical framework

Step	Question
1	Which legal provision governs the specific issue?
2	Does it directly or indirectly refer to the accounting regulatory framework?
3	Which part of the accounting regulatory framework is relevant?
4	Does its application require professional accounting judgment?
5	If so, how was professional judgment exercised and what was its result?
6	What significance does the result of professional judgment have for the application of the legal norm?

Source: Author.

The first three questions connect the legal norm with the relevant accounting framework. The fourth is the model's decision point. The fifth applies only when judgment is required, and the sixth returns the analysis to its legal purpose. Where competing professional assessments exist, the fifth step may be applied comparatively to examine the regulatory basis, data, methods, assumptions, consistency, and documentation supporting each conclusion. Accordingly, forensic accounting analysis is not required in every case in which law uses accounting categories; it becomes relevant when professional accounting judgment can affect the legally significant accounting result.

2.5. Ex ante and ex post application

The framework may be used *ex ante*, before a legal, tax, administrative, or other decision, or *ex post*, after an accounting result or legal decision exists. *Ex ante* use identifies whether a provision depends on accounting regulation and professional judgment before the accounting result acquires legal significance, while *ex post* use reconstructs the judgment underlying an existing accounting result or legal decision. The analytical sequence remains unchanged. The following hypothetical example illustrates how the six-step framework may operate *ex ante*.

Assume that, before the financial statements and the corporate income tax return are finalized, an entity holds a material trade receivable that has remained unpaid after its contractual due date and for which there are indications of increased collection risk. The entity must determine whether impairment should be recognized for accounting purposes and subsequently assess the tax consequences of that accounting treatment. The issue therefore involves a sequence extending from the legal norm, through the accounting regulatory framework and the entity's accounting policies, to professional accounting judgment and, ultimately, the legally relevant result.

Step 1 - Which legal provision governs the specific issue?

Article 16 of the Serbian Corporate Income Tax Act governs, *inter alia*, the tax treatment of write-offs and impairment of individual receivables. It connects the tax treatment of receivables with their recognition under accounting regulations and the applicable IAS/IFRS or IFRS for SMEs, while also prescribing specific conditions for the tax recognition of expenses arising from individual receivables (Corporate Income Tax Act, 2001, Art. 16). The legal analysis therefore begins with a tax provision whose application is connected with an accounting treatment but which also establishes separate conditions for tax purposes.

Step 2 - Is there a normative reference?

Yes. The reference is indirect because the tax consequence depends on an accounting treatment determined under the applicable accounting regulatory framework. Article 16 therefore establishes a connection between the tax treatment of receivables and the accounting regulatory framework. However, the accounting and tax consequences must be distinguished. Recognition of impairment as an accounting expense does not, by itself, determine its tax deductibility. The accounting treatment must first be established under the applicable accounting framework, after which the separate conditions prescribed by tax legislation must be examined (Corporate Income Tax Act, 2001, Art. 16).

Step 3 - Which accounting regulatory framework is relevant?

For an entity applying full IFRS, the relevant framework includes IFRS 9 Financial Instruments, which governs the impairment of financial assets through the expected credit loss model. In the case of trade receivables, the assessment may include the ageing of receivables, historical collection experience, information concerning the debtor's financial condition, current circumstances, and reasonable and supportable forward-looking information relevant to expected recoverability (IFRS Foundation, 2021e).

The accounting standard is, however, only one level of the analysis. The entity's accounting policies operationalize the applicable accounting requirements and define how impairment is assessed in practice. Such policies may establish ageing categories, indicators of increased collection risk, the treatment of individually significant or disputed receivables, the use of historical collection data, and other criteria relevant to recoverability. Internal procedures should further determine who performs the assessment, which documentation is required, who reviews and approves the conclusion, and how the resulting accounting adjustment is recorded.

Step 4 - Does application of the relevant accounting framework require professional accounting judgment?

Yes. Although the age of a receivable may constitute an important indicator, the assessment should not necessarily be reduced to a mechanical application of predetermined ageing percentages. Professional judgment may be required in

assessing the debtor's financial position, historical payment behaviour, subsequent collections, available collateral or guarantees, disputes concerning the receivable, current and expected economic circumstances, and other information relevant to expected recoverability.

Professional judgment is therefore exercised within the applicable accounting standard, the entity's accounting policies, and established procedures. A departure from those policies or inconsistent treatment of economically similar receivables requires particular attention and appropriate justification.

Step 5 - If so, how was professional judgment exercised and what was its result?

In an ex ante application, this question is applied prospectively. In an ex ante forensic accounting analysis, the examination is not limited to recalculating the impairment amount. It should reconstruct the decision-making process: which information was available at the assessment date, which methodology was prescribed by the accounting policies, whether that methodology was consistently applied, which assumptions were used, what evidence supported them, who performed the assessment, who approved it, and how the decision was recorded.

Particular importance should be attached to documentary evidence, including ageing schedules, contracts and invoices, correspondence with the debtor, evidence of collection attempts, subsequent payments, information on the debtor's financial position, collateral or guarantees, legal proceedings, and other evidence relevant to recoverability.

The result should be a documented conclusion as to whether the proposed impairment is supported by the applicable accounting framework, accounting policies, procedures, and available evidence. The forensic accounting function is not to replace management's estimate with an arbitrary alternative estimate, but to determine whether the professional judgment underlying the accounting result is traceable, consistently applied, and supported by evidence.

Step 6 - What significance does the result have for the application of the legal norm?

Recognition of impairment affects more than the carrying amount of the receivable. It affects expenses and consequently accounting profit and equity and may also affect financial ratios and other amounts derived from the financial statements. Where legislation assigns significance to those accounting amounts, the consequences may extend into the legal or tax sphere.

For corporate income tax purposes, however, the accounting and tax consequences must be analysed separately. An impairment recognized as an accounting expense does not automatically constitute a tax-deductible expense. Compliance with the requirements of Article 16 must be independently established. The framework therefore prevents an accounting conclusion from being mechanically transferred

into the tax determination without examining the additional conditions imposed by the legal norm.

The forensic dimension is particularly important because impairment involves estimates and professional judgment and may therefore create opportunities for intentional or unintentional bias. An unsupported or selectively applied impairment may reduce reported profit, whereas an unjustified failure to recognize impairment may have the opposite effect. This does not mean that the existence of professional judgment indicates manipulation. Rather, it identifies an area in which the risk of bias or manipulation requires stronger procedural and evidential safeguards.

Accordingly, an *ex ante* application should place particular emphasis on predefined accounting policies, consistent treatment of comparable receivables, supporting documentation, appropriate authorization, traceability of assumptions, and a clear audit trail from the underlying evidence to the accounting entry and its subsequent legal or tax treatment. Any departure from established policies or procedures should be identified, documented, and justified.

This example demonstrates the preventive function of the proposed framework. In an *ex ante* application, forensic accounting analysis is performed before the accounting result acquires its final legal or tax significance. The framework may therefore operate not only as a method for interpreting the relationship between accounting and law but also as a preventive control mechanism: it identifies where professional judgment enters the decision-making process, what evidence supports it, whether established policies and procedures have been followed, and what consequences the resulting accounting decision may produce. The analytical sequence remains the same in *ex ante* and *ex post* applications; what changes is the point in time at which it is applied.

The following section applies the framework *ex post* to judgment 9 U 10493/18 and examines its transferability through additional judicial decisions.

3. Application and testing of the analytical framework on case law

3.1. Methodological approach to the case study

The practical applicability of the framework is examined through selected decisions of Serbian courts. Administrative Court judgment 9 U 10493/18, concerning Article 7 of the Property Tax Act, is the central case study because it permits all elements of the model to be followed: the legal norm, normative reference, the relevant accounting framework, professional accounting judgment, and the legal significance of its result. Additional decisions are used only to test transferability. The purpose is not to re-decide the disputes, but to determine whether the framework systematically identifies professional judgment that may be relevant to the legal outcome.

The additional decisions were selected purposively to examine the framework across different types of accounting and valuation issues rather than to reproduce the factual

setting of the central case. The selection includes cases concerning inventory procedures, the distinction between revenue and profit, the timing of revenue recognition, and the selection of valuation methods. These cases were chosen because the legal determination depended, to different degrees, on the identification or application of an accounting or professional framework, thereby allowing the analytical sequence to be examined outside the specific context of Article 7 and fair value measurement. The selection is therefore illustrative rather than statistically representative, and no claim is made that it constitutes an exhaustive sample of Serbian case law.

3.2. Central case study - Article 7 of the Property Tax Act and judgment 9 U 10493/18

Administrative Court judgment 9 U 10493/18 concerns the application of Article 7 of the Property Tax Act and the tax treatment of the value of real estate recognised in the taxpayer's accounting records (Administrative Court of the Republic of Serbia, 2022). The analysis focuses not only on the outcome of the dispute but on the relationship between the legal norm, accounting regulation, and the professional accounting judgment that may precede determination of the value to which the law assigns legal significance.

Step 1 - Which legal provision governs the specific issue?

Article 7 of the Property Tax Act is the starting point. For certain taxpayers that keep accounting records, the provision links the tax base, subject to statutory conditions, to the fair value of real estate recognised in accordance with IAS/IFRS and adopted accounting policies (Property Tax Act, 2001, Art. 7). The statute therefore gives legal significance to an accounting value determined under the regulatory framework to which it refers.

Step 2 - Is there a normative reference?

Yes. Article 7 expressly links the tax treatment to IAS/IFRS and the taxpayer's accounting policies. For this model the reference is classified as indirect because the legal consequence depends on a previously determined accounting fair value. The analysis must therefore identify not only the amount in the accounting records but also the framework governing the determination and maintenance of that value (Property Tax Act, 2001, Art. 7).

Step 3 - Which accounting regulatory framework is relevant?

IAS 16 and IFRS 13 - Fair Value Measurement perform different but related functions. IAS 16 governs the revaluation model and the frequency of revaluations of property, plant and equipment; IFRS 13 defines fair value and establishes a framework for its measurement when measurement is required. Of particular importance is IAS 16 paragraph 34, under which the frequency of revaluations depends on changes in fair value. The Standard does not prescribe a uniform period

after which an earlier valuation automatically becomes inappropriate (IFRS Foundation, 2021b, para. 34).

This distinction separates two questions: how much time has elapsed since the previous valuation, and whether fair value has changed to an extent that requires a new valuation. The first is a factual question; the second requires the relevant accounting rule to be applied to the facts and circumstances of the case.

Step 4 - Does application of the relevant accounting framework require professional accounting judgment?

Yes. IAS 16 does not allow the need for revaluation to be determined solely by the number of years elapsed since the previous valuation. It requires an assessment of whether changes in fair value are sufficiently significant to require a new valuation (IFRS Foundation, 2021b, para. 34). The judgment may therefore require consideration of the characteristics of the property, market developments, available data, materiality of changes, the previous valuation, the accounting policy, and other relevant facts and circumstances.

Step 5 - How was professional judgment exercised and what was its result?

The forensic accounting analysis then examines the basis of the professional conclusion: when the previous valuation was performed; which accounting policy was applied; what information on changes in fair value was available; whether indicators of significant change existed; which data and assumptions supported the decision to retain or change the value; and whether the judgment was appropriately documented. The outcome is not predetermined. If available evidence indicates no significant change in fair value, judgment may support the conclusion that a new valuation was unnecessary; if it indicates a significant change, revaluation may have been required (IFRS Foundation, 2021b, para. 34).

Step 6 - What significance does the result have for the application of the legal norm?

The final step reconnects accounting analysis with Article 7. If the value in the accounting records was determined and maintained in accordance with the relevant IAS/IFRS and accounting policies, the legal significance of that fact must be assessed. If the accounting value was not based on proper application of the relevant framework, that finding may likewise be relevant. The model therefore does not favour either the taxpayer or the tax authority; it identifies the professional accounting issue that lies between the legal norm and the final legal conclusion.

Table 3 summarises the application of all six steps to the central case study.

Table 3: Application of the analytical framework to judgment 9 U 10493/18

Model step	Application to 9 U 10493/18
1. Legal norm	Article 7 of the Property Tax Act
2. Normative reference	Indirect, with an explicit link to IAS/IFRS and accounting policies
3. Accounting framework	IAS 16, IFRS 13 - Fair Value Measurement, relevant fair value rules, and the taxpayer's accounting policy
4. Professional judgment	YES - assessment of whether fair value differs materially from carrying amount and requires a new revaluation
5. Analysis of judgment	Previous valuation, accounting policy, relevant data, assumptions, valuation inputs/methods where applicable, and documentation
6. Legal significance	Assessment of whether fair value recognised in the accounting records satisfies the conditions to which Article 7 assigns legal significance

Source: Author, based on an analysis of Administrative Court judgment 9 U 10493/18, the Property Tax Act, IAS 16 and IFRS 13.

3.3. Result of Applying the Model to the Central Case Study

Application of the model shows that the relevance of the carrying amount cannot be reduced solely to the date of the previous valuation. The passage of time is a relevant fact, but by itself it does not answer whether the amount recognised in the accounting records remained in accordance with the relevant accounting regulatory framework. The key professional question is whether the circumstances and fair value changed to an extent that, under IAS 16, required a new valuation.

At the same time, the analysis points to the importance of the taxpayer's accounting policies. Article 7 of the Property Tax Act links the legally relevant fair value not only to the application of IAS/IFRS, but also to the adopted accounting policies. Therefore, excluding those policies from the analysis means that one of the elements to which the statutory provision expressly refers was not considered when assessing the appropriateness of the reported value.

From the perspective of the proposed analytical framework, accounting policies form part of the regulatory and documentary basis of professional accounting judgment. Their analysis is important for determining how the taxpayer applied the selected valuation model and whether the reported value was determined and maintained in accordance with the accounting policies and the relevant IAS/IFRS.

The proposed model therefore identifies two related methodological problems: reducing the assessment of the need for a new valuation to the passage of time since the previous valuation, and excluding accounting policies from the analysis. The forensic accounting approach allows the analysis to move beyond the final carrying amount and include the relevant standards, accounting policies, available data, assumptions, methods, and documentation on which that value was based.

The model does not predetermine the legal outcome; it identifies the accounting issues that should be analysed before a legal conclusion is reached.

3.4. Testing the transferability of the model on other court decisions

To test whether the framework extends beyond Article 7 and fair value measurement, the same six-step analytical sequence was applied to additional judicial decisions involving different accounting or valuation issues. Judgment Prev 549/2020 is presented below in greater detail as an additional structured application of the framework, while the remaining decisions are summarized in Table 4.

Step 1 - Which legal provision governs the specific issue?

The dispute concerned compensation for lost profit claimed by a shareholder on the basis of the profit that he argued could have been distributed to him. The legal determination of lost profit therefore depended, in the circumstances of the case, on establishing the amount of profit actually generated by the company.

Step 2 - Is there a normative reference?

The connection with the accounting regulatory framework is indirect. The legal rules governing compensation for lost profit do not themselves prescribe how corporate profit is calculated, but their application in this case depended on a prior determination of an accounting category – profit.

Step 3 - Which accounting regulatory framework is relevant?

The Supreme Court of Cassation referred to the accounting legislation and accounting standards governing the recognition and measurement of revenue and expenses and the determination of financial results. The relevant framework distinguishes revenue from profit: profit cannot be determined merely from invoiced or collected amounts without considering the related expenses.

Step 4 - Does application of the relevant accounting framework require professional accounting judgment?

The distinction between revenue and profit is itself a basic accounting rule and does not necessarily require complex professional judgment. However, determining the actual profit arising from transactions not recorded in the accounting records may require professional analysis and judgment concerning the nature of transactions, related revenues and expenses, recognition periods, available documentation, and the reliability of the underlying data.

Step 5 - If so, how was professional judgment exercised and what was its result?

The analysis requires reconstruction of the transactions underlying the unrecorded receipts, including invoices, the basis of payments, related costs, recognition periods, bank documentation, and the reliability of the available evidence. Only after revenues and the corresponding expenses have been properly identified can the relevant financial result be determined. The judgment itself illustrates the importance of this step because the expert evidence relied partly on subsequently submitted documentation whose reliability and completeness were disputed.

Step 6 - What significance does the result have for the application of the legal norm?

The accounting result directly affects the legal determination of lost profit. If revenue is incorrectly treated as profit, the accounting error is transferred to the legal assessment of damages. The Supreme Court of Cassation therefore found that the relevant profit had not been properly established and that the lower courts had effectively equated the shareholder's potential profit with the company's revenue in the disputed period (Supreme Court of Cassation, 2021).

Table 4: Testing the applicability of the analytical framework

Court decision	Professional/accounting issue	Problem relevant to the model	Result
9 U 10493/18	Fair value of real estate and the need for revaluation	Assessment of whether changes in fair value require a new valuation under the relevant accounting framework	Central case study - model applicable
Rev2 3223/2020	Inventory count and distinction between accounting inventory and operational stock control	Need to identify the nature of the inventory process correctly before applying rules on inventory of assets and liabilities	Illustrative application
Prev 549/2020	Distinction between revenue, expenses and profit	Determining profit requires a correct understanding of the accounting result, not merely revenue earned	Detailed transferability test – model applicable
Pkž 293/18	Timing of revenue recognition and application of IFRS for SMEs	Identification of the relevant accounting rule and correctness of timing of revenue recognition	Illustrative application
Prev 862/2022 / Prev 1018/2024	Selection of a share valuation method	Selection of an appropriate valuation method requires professional justification based on professional rules and case-specific circumstances	Illustrative application

Source: Author.

The structured application to Prev 549/2020 demonstrates that the analytical sequence can be applied beyond the central fair value case and to a different relationship between accounting and law. The remaining decisions provide additional illustrations of the framework's applicability to different accounting and valuation issues. Importantly, the framework does not assume that every accounting issue necessarily requires professional judgment; rather, it identifies the point at which professional assessment becomes relevant and distinguishes it from the mechanical application of an accounting rule.

Conclusion

This research addresses situations in which the application of a legal norm depends on an accounting category whose content derives from the accounting regulatory framework, particularly where the relevant accounting rule requires professional accounting judgment. In such cases, proper understanding of the legally relevant accounting result may require analysis not only of the final amount or accounting treatment, but also of the professional process through which that result was reached.

The proposed methodological framework identifies the legal norm, determines whether it directly or indirectly refers to the accounting regulatory framework, identifies the relevant accounting rule, and asks whether applying that rule requires professional accounting judgment. If judgment is not required, the relevant accounting rule may be applied directly. If it is required, the forensic accounting analysis examines how professional judgment was exercised, the information, accounting policies, methods, assumptions and documentation on which it was based, its resulting accounting effect, and the significance of that result for the application of the legal norm.

Application of the framework to Administrative Court judgment 9 U 10493/18 demonstrates its practical operation in an ex post setting. The analysis shows that the need for revaluation cannot be assessed solely by the passage of time and that accounting policies cannot be excluded where Article 7 of the Property Tax Act expressly refers to them. The framework therefore identifies the accounting issues that should be examined before a legal conclusion is drawn, without predetermining the legal outcome.

The additional structured application to judgment Prev 549/2020 demonstrates that the same analytical sequence can be applied beyond the central fair value case and to a different relationship between accounting and law. The remaining selected judicial decisions provide further illustrations of the framework's applicability to different accounting and valuation issues. These applications also show that the framework does not presume that every accounting or expert issue necessarily requires professional accounting judgment; rather, it identifies the point at which such judgment becomes relevant and distinguishes it from the mechanical application of an accounting rule.

The hypothetical ex ante application further demonstrates the preventive dimension of the framework. By examining impairment of trade receivables before the accounting result acquires its final legal or tax significance, the framework identifies where professional judgment enters the decision-making process, which accounting policies and procedures govern that judgment, what evidence supports it, and what accounting, legal, or tax consequences may follow. Ex ante and ex post applications therefore follow the same analytical sequence; the principal difference is the point in time at which the analysis is performed.

The scientific contribution of the research is a methodological framework that systematically connects the legal norm, the accounting regulatory framework, professional accounting judgment, and the legal significance of its result. The forensic accounting perspective adds a structured reconstruction of the decision process through analysis of applicable rules and accounting policies, available facts and data, methods, assumptions, supporting documentation, and the consistency of the resulting professional conclusion. In this way, forensic accounting can serve not only a reconstructive function after a dispute has arisen, but also a preventive function before an accounting judgment produces legally relevant consequences.

The framework may assist courts, tax and other authorities, lawyers, accountants, auditors, and forensic accountants when an accounting result has legal significance. Its purpose is neither to replace legal decision-making nor to substitute an external judgment for professional accounting judgment, but to provide a structured methodological connection between them. The principal limitation of the study is the small and purposively selected set of judicial decisions analysed. Further research should therefore examine the framework on a broader and more diverse sample of judicial, tax, and administrative cases, including other areas in which legal consequences depend on accounting estimates, classifications, valuations, or other forms of professional accounting judgment.

References

- Ajekwe, C. C. (2022). Application of professional judgement in International Financial Reporting Standards. *European Journal of Accounting, Auditing and Finance Research*, 10(7), 17–30. <https://doi.org/10.37745/ejaaf.2013/vo10.n7pp1730>
- Bennett, B., Bradbury, M., & Prangnell, H. (2006). Rules, principles and judgments in accounting standards. *Abacus*, 42(2), 189–204. <https://doi.org/10.1111/j.1467-6281.2006.00197.x>
- Bonner, S. E. (1999). Judgment and decision-making research in accounting. *Accounting Horizons*, 13(4), 385–398. <https://doi.org/10.2308/acch.1999.13.4.385>
- Cannon, N. H., & Bedard, J. C. (2017). Auditing challenging fair value measurements: Evidence from the field. *The Accounting Review*, 92(4), 81–114. <https://doi.org/10.2308/accr-51569>

Fomina, O., Zadniprotsky, O., Korol, S., & Romashko, O. (2022). Professional judgement in accounting: Contents and conditions of application. *Business: Theory and Practice*, 23(1), 26–38. <https://doi.org/10.3846/btp.2022.13330>

Griffith, E. E., Hammersley, J. S., Kadous, K., & Young, D. (2015). Auditor mindsets and audits of complex estimates. *Journal of Accounting Research*, 53(1), 49–77. <https://doi.org/10.1111/1475-679X.12066>

Honigsberg, C. (2020). Forensic accounting. *Annual Review of Law and Social Science*, 16, 147–164. <https://doi.org/10.1146/annurev-lawsocsci-020320-022159>

IFRS Foundation. (2021a). IAS 8: Accounting Policies, Changes in Accounting Estimates and Errors. IFRS Foundation.

IFRS Foundation. (2021b). IAS 16: Property, Plant and Equipment. IFRS Foundation.

IFRS Foundation. (2021c). IAS 36: Impairment of Assets. IFRS Foundation.

IFRS Foundation. (2021d). IFRS 13: Fair Value Measurement. IFRS Foundation.

IFRS Foundation. (2021e). IFRS 9: Financial Instruments. IFRS Foundation.

International Auditing and Assurance Standards Board (IAASB). (2018). *ISA 540 (Revised), Auditing Accounting Estimates and Related Disclosures*. International Federation of Accountants.

Krpić, J. (2023). Profesionalno prosuđivanje u kontekstu međunarodnih standarda finansijskog izveštavanja. *Revizor*, 26(101), 71–79. <https://doi.org/10.56362/Rev23101071K>

Mala, R., & Chand, P. (2015). Judgment and decision-making research in auditing and accounting: Future research implications of person, task, and environment perspective. *Accounting Perspectives*, 14(1), 1–50. <https://doi.org/10.1111/1911-3838.12040>

Tekavčić, M., & Damijan, S. (2021). Forensic accounting vs fraud examination: Roles, importance and differences. *The Journal of Forensic Accounting Profession*, 1(2), 29–47. <https://doi.org/10.2478/jfap-2021-0007>

Trotman, K. T. (1998). Audit judgment research—Issues addressed, research methods and future directions. *Accounting & Finance*, 38(2), 115–156. <https://doi.org/10.1111/1467-629X.00007>

Trotman, K. T., Tan, H. C., & Ang, N. (2011). Fifty-year overview of judgment and decision-making research in accounting. *Accounting & Finance*, 51(1), 278–360. <https://doi.org/10.1111/j.1467-629X.2010.00398.x>

Zakon o bankama [Banking Act]. (2005). Službeni glasnik Republike Srbije.

Zakon o porezima na imovinu [Property Tax Act]. (2001). Službeni glasnik Republike Srbije.

Zakon o privrednim društvima [Companies Act]. (2011). Službeni glasnik Republike Srbije.

Zakon o računovodstvu [Accounting Act]. (2019). Službeni glasnik Republike Srbije.

Zakon o tržištu kapitala [Capital Market Act]. (2021). Službeni glasnik Republike Srbije.

Zakon o stečaju [Bankruptcy Act]. (2009). Službeni glasnik Republike Srbije.

Zakon o porezu na dobit pravnih lica [Corporate Income Tax Act]. (2001). Službeni glasnik Republike Srbije.

Upravni sud Republike Srbije [Administrative Court of the Republic of Serbia]. (2022, February 3). Judgment 9 U 10493/18.

Privredni apelacioni sud [Commercial Appellate Court]. (2018). Pkž 293/18. <https://sudskapraksa.rs/dokument/pkz-293-2018-presuda-kojom-se-potvrđuje-osudjujuca-presuda-za-vise-privrednih-prestupa-iz-racunovodstva-c15fe86a27>

Vrhovni kasacioni sud [Supreme Court of Cassation]. (2021, March 4). Prev 549/2020: Lost profit. <https://vrh.sud.rs/sr-lat/prev-5492020-312832-izmakla-korist>

Vrhovni kasacioni sud [Supreme Court of Cassation]. (2021, February 24). Rev2 3223/2020. <https://www.vrh.sud.rs/sr-lat/rev2-32232020-35163>

Vrhovni kasacioni sud [Supreme Court of Cassation]. (2022, December 15). Prev 862/2022: Examination of the adequacy of share price. <https://vrh.sud.rs/sr-lat/prev-8622022-ispitivanje-primerenosti-cene-akcije>

Vrhovni sud [Supreme Court]. (2025, February 27). Prev 1018/2024. <https://www.vrh.sud.rs/sr-lat/prev-10182024-32173>

Borko Nikolić
Attorney at Law

INSTITUTIONAL MODELS OF FORENSIC ACCOUNTING EXPERTISE IN FINANCIAL CRIME PROCEEDINGS: COMPARATIVE PERSPECTIVES AND IMPLICATIONS FOR SERBIAN LAW

UDK 343.1:657.6 (497.11)

DOI <https://doi.org/10.54318/czf.2026.bn.027>

Abstract

The growing complexity of financial crime has increased the demand for specialised forensic accounting expertise capable of supporting investigation and adjudication of complex financial crime. Comparative legal systems have institutionalised such expertise through different procedural arrangements, most commonly through court appointed experts, party commissioned forensic accountants, and specialised forensic accounting services embedded within public prosecution offices. While each arrangement pursues the common objective of improving the quality of financial evidence, they differ in terms of institutional independence, procedural function, evidentiary status, and their interaction within criminal proceedings. The purpose of this paper is to identify the principal procedural functions of forensic accounting expertise, examine the solutions adopted in selected comparative jurisdictions, and assess whether the Serbian hybrid framework provides a sufficiently coherent procedural relationship between investigative, court oriented, and party based forms of forensic accounting expertise.

The paper argues that the effectiveness of forensic accounting depends less on the predominance of any particular institutional arrangement than on the existence of clear procedural rules governing the respective functions, interaction, and judicial evaluation of different forms of forensic accounting expertise. Through a functional comparative analysis, the paper identifies common procedural principles governing the allocation of forensic accounting expertise among investigative authorities, parties, and the courts. Although legal systems institutionalise these functions differently, they place emphasis on clearly defined procedural roles, methodological reliability, effective judicial scrutiny, and safeguards ensuring procedural fairness.

Applying these comparative principles to Serbia, the paper examines whether the interaction between court appointed experts, party commissioned forensic accounting expertise, and the specialised Financial Forensic Service within the Public Prosecutor's Office is sufficiently articulated by the existing procedural framework. It evaluates the extent to which legislation and emerging judicial

practice provide clear guidance regarding their respective procedural functions, evidentiary relationship, and judicial evaluation. Against this comparative backdrop, the paper proposes legislative and institutional measures aimed at strengthening procedural coherence, professional standards, and the effective use of forensic accounting expertise in complex financial crime proceedings.

Key words: *forensic accounting, expert evidence, financial crime, criminal procedure, comparative law, institutional models, Serbia*

JEL Classification: *K41, K42.*

Introduction

The growing complexity of financial crime has strengthened the need for specialised financial expertise in criminal proceedings, while also increasing the demands placed on prosecutors and judges dealing with such cases. Investigations and prosecutions involving fraud, corruption, money laundering and related offences frequently require the analysis of complex financial records, reconstruction of transactions, identification of financial relationships and assessment of the economic consequences of criminal conduct. Effective handling of such cases therefore depends not only on the development of relevant professional capacities within the judiciary and prosecution, but also on appropriate modalities for accessing specialised financial expertise. Within these modalities, forensic accounting may perform functions that extend beyond the provision of technical accounting knowledge, particularly where financial analysis forms an important part of establishing the factual basis of a criminal case.

The institutional position of forensic accounting expertise, however, differs considerably across legal systems. Depending on the procedural framework, financial expertise may be produced by experts appointed by the court, commissioned by parties to the proceedings, or carried out within specialised investigative or prosecutorial structures. These arrangements are not merely organisational alternatives. They determine who initiates and controls the examination, the procedural status of the person providing the expertise, the relationship between financial analysis and the collection of evidence, and the manner in which the relevant findings may be challenged and evaluated in criminal proceedings.

This paper examines selected institutional models from a comparative perspective, with particular emphasis on their procedural implications in criminal proceedings concerning financial crime. The analysis is focused on the criminal justice context. It does not seek to examine the use of forensic accounting in civil litigation, arbitration, tax proceedings, internal corporate investigations or other non criminal contexts, except where a reference is necessary to clarify a particular institutional model. This delimitation reflects the specific evidentiary requirements and procedural safeguards of criminal proceedings, where the production and evaluation

of evidence must take into account the rights of the accused, the role of investigative and prosecutorial authorities, and fair trial standards.

The research adopts a comparative legal methodology combining doctrinal and functional approach. Firstly, the relevant procedural and institutional arrangements in selected jurisdictions are examined on the basis of their legal framework and the role assigned to forensic accounting expertise within criminal proceedings. Secondly, the selected models are compared according to their principal procedural characteristics. Particular attention is given to the relationship between investigative authorities, courts, parties and forensic accounting experts, the procedural stage at which financial expertise is introduced, and the mechanisms governing the production, presentation and evaluation of expert findings. Finally, the comparative findings are used as a basis for assessing the coherence of the Serbian framework and, in particular, the relationship between investigative, court oriented and party based forms of forensic accounting expertise.

The comparative analysis focuses on selected jurisdictions representing different approaches to the institutionalisation of forensic accounting expertise. The purpose is not to provide an exhaustive survey of national systems, but to identify distinct institutional solutions and the procedural functions associated with them that can serve as analytical reference points for assessing the Serbian legal framework.

Against this backdrop, the paper proceeds in three principal stages. The first part identifies the principal procedural functions of forensic accounting expertise in financial crime proceedings and establishes the analytical framework for the comparison. The second part examines selected institutional models, with particular reference to the United States and Germany, focusing on the respective roles of court appointed experts, party commissioned experts and specialised forensic accounting services operating within prosecutorial or investigative structures. The third part applies the comparative findings to Serbian law and considers whether its hybrid framework establishes a sufficiently coherent procedural relationship between these different forms of forensic accounting expertise.

The central argument is that the institutional organisation of forensic accounting expertise may differ substantially between legal systems, but that the institutional model alone does not determine the procedural reliability of financial expertise. Irrespective of whether expertise is situated within the court, the parties or investigative and prosecutorial structures, its effective use in criminal proceedings ultimately depends on procedural rules governing its production, independence, scrutiny and evaluation. The Serbian framework is therefore assessed not by reference to a single preferred institutional model, but by examining whether its combination of different forms of forensic accounting expertise provides a coherent and procedurally adequate framework for financial crime proceedings.

1. Comparative Institutional Arrangements

1.1. Party Initiated Forensic Accounting Expertise and Procedural Safeguards

In the federal criminal justice system of the United States, expert testimony is governed principally by Federal Rule of Evidence 702, supplemented by Rules 703 and 705 concerning the bases for expert opinions and disclosure of the underlying facts or data at trial. The Federal Rules of Criminal Procedure separately regulate the pretrial disclosure of expert testimony by the prosecution and the defence through Rule 16(a)(1)(G) and Rule 16(b)(1)(C). The research conducted by Singleton & Singleton specifically addresses the role of the forensic accountant as a witness in court, legal qualifications, admissibility of accounting evidence, and the forensic accountant's role in the litigation team (Singleton A., Singledon T., 2010). According to this study the established framework permits both parties to present expert evidence where the requirements governing expert testimony are satisfied. In financial crime proceedings, this framework can accommodate specialised expertise concerning accounting records, financial transactions, financial losses, asset movements and other matters requiring specialised knowledge. The institutional characteristic relevant for this comparison is that such expertise operates within an adversarial evidentiary framework in which the parties may present their own expert evidence, subject to judicial control over admissibility, including an assessment of whether the proposed expert opinion satisfies the applicable requirements of reliability and relevance.

In *Daubert* case, the U.S. Supreme Court held that Rule 702 requires trial judges to act as 'gatekeepers' and ensure that expert testimony is both relevant and reliable. The Supreme Court identified several factors that may assist in assessing reliability, including: whether a theory or technique can be and has been tested, whether it has been subjected to peer review and publication, its known or potential error rate, the existence of standards controlling its operation, and its acceptance within the relevant professional or scientific community. The Court emphasised, however, that these factors do not constitute a definitive checklist and that the inquiry must remain flexible and connected to the particular circumstances of the case.

The U.S. Supreme Court subsequently clarified in *Kumho Tire Co.* that this 'gatekeeping' obligation applies not only to scientific testimony but also to technical and other specialised expert testimony falling within Rule 702. The specific *Daubert* factors therefore need not apply mechanically to every form of expertise. Rather, the trial court must determine which considerations provide appropriate measures of reliability in the particular case. This is particularly relevant to forensic accounting expertise, which may involve specialised technical knowledge rather than scientific expertise in the narrower sense.

This approach is significant because it addresses the risk of disproportionate reliance on expert evidence by strengthening the procedural safeguards through which its competence, methodology and conclusions can be scrutinised. Party initiated

expertise is not treated as inherently neutral merely because the expert possesses specialised knowledge. Its evidentiary significance depends on the expert's qualifications, the reliability of the methodology and factual basis of the opinion, while disclosure to the opposing party and the opportunity for adversarial examination provide mechanisms through which the evidence may be tested. The court ultimately retains responsibility for determining whether the proposed expert evidence may be admitted and for evaluating its significance within the evidentiary record.

The principal comparative value of this approach lies in the distinction between the origin of the expertise and its procedural evaluation. The fact that specialised analysis has been commissioned by a party does not necessarily prevent it from participating in the evidentiary process. At the same time, party commissioning does not automatically confer evidentiary authority on the resulting report. Rather, procedural rules provide a framework through which the expertise may be disclosed, tested and subjected to judicial scrutiny.

This distinction is also relevant to the Serbian system, where the institution of professional consultant (*stručni savetnik*) enables a party to engage a person with specialised knowledge in relation to an ordered expert examination, thereby providing a procedural mechanism through which party associated expertise can participate in, scrutinise and challenge formally appointed expert evidence without itself acquiring the status of the expert appointed by the court or public prosecution.

1.2. Court Appointed Forensic Accounting Expertise and Judicial Control

The German criminal justice system provides a contrasting institutional model in which expert evidence is principally integrated into the proceedings through judicial appointment. Under Section 73 of the German Code of Criminal Procedure (*Strafprozessordnung, StPO*), the selection and number of experts to be consulted are determined by the judge. Where experts are publicly appointed for particular types of opinions, the provision further indicates a preference for their selection unless special circumstances justify the appointment of others. This judicially controlled model is particularly characteristic of the trial stage. During the investigation, the public prosecution office and police may also engage experts of their choice. The parties in the proceedings do not therefore initiate the appointment of the expert in the same manner as under the party oriented model described above, although they retain procedural rights concerning the appointed expert, including the right to challenge an expert under Section 74 StPO.

The institutional position of the expert is further defined by judicial control over the expert examination. Section 78 StPO provides that the judge directs the expert's participation in the proceedings where considered necessary. Section 80 StPO permits the expert, for purposes of preparing the opinion, to obtain further clarification and, subject to the applicable conditions, inspect the case files and attend examinations. The court may also order a new expert opinion where it

considers the existing opinion insufficient under Section 83 StPO. Unlike in the previous model, the resulting framework therefore places the expert within an evidentiary structure that is judicially controlled rather than treating expert analysis primarily as evidence generated by the parties.

Section 244(2) StPO places the court under an *ex officio* duty to investigate relevant facts. Section 244(4) StPO allows the court to refuse a further expert opinion in certain circumstances. It also provides for a further expert where there are doubts about the competence of the existing expert, incorrect factual assumptions, contradictions in the opinion, or better research methods that could produce more reliable results.

For forensic accounting and other forms of financial expertise, the significance of this model lies in the institutional separation between the authority commissioning the expertise and the parties whose interests may be affected by its conclusions. A court appointed expert is not appointed for either side but performs a specialised evidentiary function within proceedings directed by the court. This institutional position may strengthen perceptions of impartiality and reduce the risk that expert analysis is shaped by the interests of the party commissioning it. At the same time, judicial control does not eliminate the need for procedural safeguards. The parties retain rights to challenge an appointed expert, and the court may hear further expert evidence where the conditions for doing so are satisfied.

The German criminal procedure combines institutional separation from the parties at the point of appointment with procedural mechanisms through which the parties can contest the expert evidence. The effectiveness of this model nevertheless depends on the court's ability to formulate appropriate instructions, critically assess specialised evidence and ensure that the expert's technical assessment does not displace the court's own responsibility for determining the facts. The provisions of StPO itself reflect this responsibility: the court directs the expert's activity, may order a new opinion where an existing opinion is insufficient, and must assess whether further expert evidence is warranted in the case.

The German model therefore illustrates both an advantage and a potential limitation of judicially controlled expertise. Judicial appointment provides an institutional basis for separating the expert from the interests of the parties, but it also places considerable responsibility on the court to select appropriate expertise, define the scope of the examination and critically evaluate the resulting opinion. German literature has highlighted and discussed this duality in the practical operation of expert evidence since the 1970s. For example, Weigand, discussing earlier Karl Peters's analysis of complexities with reopened criminal proceedings, notes that the increasing sophistication of scientific expertise does not eliminate the possibility of erroneous conclusions and identifies deficiencies both in the preparation of expert opinions and in their judicial evaluation (Weigand, 1976). In particular, both authors indicate that courts may inadequately assess the expert's qualifications and expertise, misunderstand specialised terminology, or fail to examine conflicting expert

opinions sufficiently. Judicial control consequently operates not as a guarantee that expert evidence is correct, but as a procedural mechanism through which the expertise is commissioned, challenged and, where necessary, supplemented or replaced.

1.3 Investigative Forensic Accounting Expertise and Prosecutorial Functions

A different institutional model exists in England and Wales in which forensic accounting expertise may be integrated into the investigative and prosecutorial process rather than being initiated primarily by the court or by the parties. This model is particularly visible in functioning the United Kingdom's Serious Fraud Office (SFO), established under the 1987 Criminal Justice Act. It combines investigative and prosecutorial functions and operates through multidisciplinary case teams involving lawyers, investigators, forensic accountants and other specialists. This organisational structure enables financial analysis to be integrated into the investigation and prosecution of serious and complex economic crime.

The institutional role of the forensic accountant in this model is therefore not necessarily limited to the preparation of an opinion to be presented at trial. Financial analysis may form part of the investigative process itself. It may include the examination of accounting records, reconstruction of transactions, assessment of financial discrepancies and analysis of the financial consequences of suspected wrongdoing. A concrete illustration can be provided in the case *R v H (2007)*, in which a forensic accountant working for the SFO examined the relevant documentation, made allowances for missing documents and inadequacies in IKEA's accounting system, excluded certain transactions from consideration and reported an apparent shortfall of approximately 7 million British pounds. The prosecution relied on that analysis in the proceedings. The case illustrates how forensic accounting expertise can originate within an investigative agency and subsequently acquire a formal evidentiary function in criminal proceedings.

This distinction is important because investigative expertise should not automatically be equated with expert evidence provided by a forensic accountant. An accountant working within an investigative or prosecutorial team may contribute specialised knowledge to the analysis and development of a case without necessarily assuming the procedural position of an expert witness. However, once an opinion is relied upon as expert evidence the expert is subject to a distinct procedural framework. Under Part 19 (Rules 19.1 to 19.4) of the Criminal Procedure Rules, where a party seeks to introduce expert opinion evidence, an expert must assist the court by giving an opinion that is objective and unbiased and within the expert's area of expertise. This duty is owed to the court and overrides any obligation to the party instructing the expert. The expert must also disclose information that might reasonably be thought capable of undermining the opinion or detracting from the expert's credibility or impartiality.

The institutional implications of this model have been thoroughly examined in academic literature. In his assessment of the SFO, established following the recommendations of the 1986 Fraud Trials Committee, Levi reconsidered the institutional arguments for a unified approach to serious fraud and the subsequent development of the this model, emphasising the rationale for combining investigative and prosecutorial functions within a specialised institutional structure (Levi, 2003). Research on forensic accounting in the United Kingdom by a group of authors has, in turn, characterised the field as multidisciplinary, involving complementary legal, investigative and financial expertise rather than a narrowly defined accounting function (Hegazy, Sangster and Kotb, 2017). These perspectives are relevant to the present analysis because the SFO model incorporates forensic accountants within a broader institutional structure in which specialised financial analysis forms part of the investigative process.

The English model combines an investigative function with procedural safeguards applicable when specialised analysis is introduced as expert evidence creating a different institutional relationship from the US and German models. The forensic expert may participate in the analysis and development of the prosecution case from the investigative stage, but the resulting expert evidence does not acquire authority merely from its institutional origin. Its use in proceedings is subject to disclosure, impartiality requirements and judicial control, while the opposing party must have an opportunity to challenge the analysis and its underlying basis. This model illustrates that forensic accounting expertise does not have to be institutionally confined to a formal court or prosecutorial expert examination. It may perform a distinct investigative function within the prosecution or investigative structure, while procedural safeguards become particularly important when the resulting analysis is introduced as expert evidence.

1.4. Comparative Observations

The three models demonstrate that forensic accounting expertise may perform different institutional functions in criminal proceedings. The U.S. model emphasises party initiated expertise subject to judicial admissibility control and adversarial testing. The German model places greater responsibility for the selection and structuring of expert examination with the court, while England and Wales illustrates the integration of specialised financial expertise into the investigation and prosecution of complex economic crime. However, many authors caution against treating adversarial and inquisitorial arrangements as rigid or mutually exclusive models and reiterates that their procedural consequences depend on how expertise is generated, introduced and tested within the particular institutional setting (see for example, Roberts, 2015).

The comparative analysis therefore does not establish that one model provides inherently more reliable expertise than the other. Party initiated expertise may facilitate competing analysis and adversarial challenge, but may also raise concerns about party influence. Judicially controlled expertise may strengthen institutional

independence, but judicial appointment does not remove the need for critical assessment of the expert opinion. Investigative expertise may provide valuable specialised analysis at an early stage of financial crime proceedings, but its institutional role must be distinguished from that of formal expert evidence. As Decaigny points out in his study, the central issue with comparative analysis is consequently not the choice between adversarial and inquisitorial structures, but whether the procedural arrangements enable specialised knowledge to be generated, communicated and evaluated in a manner consistent with the requirements of a fair and reliable criminal process (Decaigny, 2014).

For the purposes of this paper, the comparison provides a framework for assessing the Serbian system without assuming that any foreign model should be adopted in its entirety. The relevant question is therefore whether Serbian law establishes a coherent relationship between investigative, court oriented and party associated forms of forensic accounting expertise, and whether the procedural safeguards applicable to each form adequately address the risks arising from its institutional position.

2. Forensic Accounting Expertise in the Serbian Legal System

The primary procedural mechanism for obtaining specialised knowledge in Serbian criminal proceedings is through expert examination. Article 113 of the Criminal Procedure Code (CPC) expressly provides that expert examination is ordered when specialised knowledge is necessary for the establishment or assessment of a fact, while excluding legal questions from its scope.

The CPC regulates the ordering and conduct of expert examination, the appointment of experts, their rights and duties, the preparation of findings and opinions, and the examination of expert evidence within criminal proceedings. In particular, Article 114 regulates the selection of the expert and establishes the order in which an expert examination is to be entrusted to a permanent court expert, a specialised institution or a state body, depending on the circumstances prescribed by the Code. Article 117 determines how expert examination is ordered. The procedural framework under the Code does not therefore limit expert examination to an individual appointed directly by a court. Rather, it provides for different institutional sources of specialised expertise and determines the order in which they are to be considered.

The institutional framework applicable to court experts in Serbia is supplemented by the Law on Court Experts, which regulates the conditions and procedures for the appointment of individual experts and the performance of expert activities by legal entities and state bodies authorised to conduct expert examinations. The Law also establishes requirements relating to professional competence and the performance of expert duties. Its purpose is therefore primarily institutional and professional, while the CPC determines the procedural conditions under which expert examination is used in criminal proceedings.

In proceedings involving complex financial crime, expert examination may be used where the determination or assessment of particular facts requires specialised knowledge in economics, finance or accounting. Depending on the factual issues to be resolved, the examination may address, for example, the analysis of financial transactions, accounting records, financial documentation, or other matters requiring specialised expertise. The findings and opinion produced through the expert examination are then introduced and examined in accordance with the procedural rules governing expert evidence and are subject to assessment by the authority conducting the proceedings taken together with the other evidence in the case.

In proceedings involving complex financial crime, expert examination may be used where the determination or assessment of particular facts requires specialised knowledge in economics, finance or accounting. Depending on the factual issues to be resolved, the examination may address, for example, financial transactions, accounting records and financial documentation. The institutional role of such expertise differs according to the procedural stage. During the investigation, expert examination may assist the public prosecutor in establishing and assessing facts relevant to the case and in deciding on further prosecutorial action. At trial, the court assumes responsibility for the conduct and assessment of the evidentiary process. The expert findings and opinion are presented and examined before the court, which evaluates them together with the other evidence. Thus, while the prosecutor may rely on expert examination as an investigative tool, the court must critically assess the expert's findings and opinion as part of the evidence and retain responsibility for determining the facts and applying the law. The expert provides specialised knowledge at both stages but does not determine criminal responsibility or replace the judicial function. This distinction is particularly relevant in financial crime proceedings, where complex accounting or financial analysis may be essential to establishing the factual basis of the case but cannot itself determine the legal consequences arising from those facts.

2.1. Professional Assistance Selected by the Parties

The CPC also recognises the possibility for a party to obtain specialised assistance in relation to expert examination. Articles 125 and 126 of the CPC regulate the role of the professional consultant (*stručni savetnik*), who may be selected and authorised by a party under the conditions provided for by the CPC.

The role of the professional consultant is procedurally related to expert examination ordered within the criminal proceedings. The consultant may participate in the examination in the manner provided by the Code, inspect relevant materials, make observations and proposals, and examine or challenge the findings and opinion of the expert. The professional consultant may also be examined in relation to matters falling within the subject of expert examination.

This legal arrangement is significant because it introduces an adversarial element into the use and evaluation of expertise. Expert examination ordered within criminal

proceedings does not operate in complete isolation from the parties. The procedural framework enables parties to obtain specialised assistance and to participate more effectively in the examination and assessment of complex expert findings.

Nevertheless, the procedural position of the professional consultant is distinct from, and should not automatically be equated with, that of a privately commissioned expert commonly found in other legal systems. The Serbian legal framework links the professional consultant to an expert examination conducted within the procedural framework of the CPC. The consultant's role is therefore primarily to provide specialised professional assistance, participate in the examination, and challenge the expert's findings and opinion.

Nevertheless, the practical implementation of this institute raises a number of questions. Serbian authors have observed that the professional consultant is in practice engaged primarily by the defence, assisting the accused in understanding expert findings, formulating objections and addressing specialised issues arising from the expert examination (Janković, 2025). Janković considers this role important for strengthening the position of the defence and ensuring the effective application of adversarial proceedings and equality of arms, and notes that the proposed amendments to the CPC would further strengthen the consultant's participation in the examination and regulate the relationship between the expert's findings and opinion and the professional consultant's expert opinion. The practical significance of the institute therefore depends not only on the formal possibility of engaging specialised assistance, but also on the consultant's ability to participate meaningfully in the examination and contribute to the critical assessment of expert evidence.

2.2. Financial forensic expertise within the Public Prosecutor's Office

Another form of specialised financial expertise was introduced through the Law on the Organisation and Jurisdiction of State Authorities in the Suppression of Organised Crime, Terrorism and Corruption. Article 19 of this Law provides for the establishment of a Financial Forensic Service within the public prosecution structures established within higher public prosecutor's office and defines the role of financial forensic experts.

Unlike formal expert examination under the CPC, the financial forensic expert is not engaged to conduct an expert examination within the procedural regime applicable to court experts. The expert instead provides specialised financial analysis to support the public prosecutor in investigating and prosecuting financial crime, including the analysis of financial flows, transactions and complex financial relationships. According to a group of Serbian authors, this institutional role enables financial analysis to form part of the investigation itself. Financial forensic experts may reconstruct transactions, analyse complex financial relationships and identify discrepancies or patterns relevant to the investigation. Their work may therefore assist the public prosecutor in identifying relevant facts and evidence before a need for formal expert examination arises (Sinanović et al., 2021).

However, the effectiveness of this institutional model is also affected by capacity constraints: the insufficient number and uneven distribution of financial forensic experts and limited operational integration of financial forensic expertise within specialised prosecutorial departments (Kolaković-Bojović & Matić-Bošković, 2025). The authors also identify gaps in the specialised knowledge and skills of public prosecutors, including the need to strengthen their capacity to conduct complex financial investigations and to make effective use of financial forensic expertise. This is significant not only because prosecutors need access to appropriate expertise, but also because they must be able to identify the type of financial expertise required in a particular case and understand and use its results in developing the investigation and prosecution.

The analytical work of a financial forensic expert may overlap with matters subsequently addressed through formal economic or financial expert examination. The Financial Forensic Service has a statutory basis for providing investigative support to the public prosecutor, but the legislation does not comprehensively define the procedural status of its analyses or reports when they become relevant to criminal proceedings. This raises an important question for the Serbian legal framework: how should investigative financial analysis relate to formal expert evidence within the same proceedings?

The issue has acquired particular significance in judicial practice. In its recent decision Kzz 621/25, the Supreme Court of Serbia addressed for the first time the procedural position of a report prepared by a financial forensic expert engaged within the prosecutorial framework. The Court accepted that such a report could be used as evidence in criminal proceedings and could form part of the evidentiary basis on which a judicial decision is made. At the same time, the Court distinguished the report of a financial forensic expert from formal economic and financial expert examination and made clear that the former does not replace the latter.

This distinction is important for understanding the Serbian institutional framework. It indicates that financial forensic expertise within the public prosecutor's office cannot simply be classified as another form of court expert evidence. Rather, it represents a specialised investigative form of expertise whose analytical product may, subject to the applicable procedural requirements and judicial evaluation, acquire evidentiary relevance without losing its distinct institutional and procedural character.

It can be argued that the Supreme Court decision helps clarify the relationship between investigative financial expertise and formal expert examination. It confirms that the existence of specialised financial knowledge does not, by itself, determine the procedural status of the resulting analysis. That status depends on the legal mechanism through which the expertise is generated and on how its results are introduced and evaluated in the criminal proceedings. The decision, however, does not fully resolve the evidentiary status of reports produced by financial forensic

experts in criminal proceedings, leaving this issue for further consideration in assessing the coherence of the Serbian legal framework.

2.3. Procedural Relationship between the Forms of Expertise

Taken together, the Serbian legal and institutional framework provides three distinct modalities of expertise through which specialised financial knowledge may enter or support criminal proceedings. Formal expert examination under the CPC produces findings and opinions within a defined procedural framework. The professional consultant provides a party with specialised assistance in examining and challenging such expertise. The Financial Forensic Service provides specialised financial analysis in support of the investigation and prosecution of financial crime.

Although these functions are institutionally distinct, they may concern the same financial facts and evidence. Financial analysis conducted during an investigation may subsequently become relevant to formal expert examination. Similarly, a professional consultant may assist a party in assessing either a formal expert opinion or financial analysis produced during the investigation. The court may therefore encounter specialised analyses generated by different actors, for different purposes and under different procedural conditions.

The issue is consequently not whether one form of expertise should replace another, but whether their respective procedural status and evidentiary significance are sufficiently clear. The Supreme Court's decision is relevant in this respect because it indicates that a report produced by a financial forensic expert may have evidentiary relevance without being equivalent to a formal expert opinion under the CPC. The decision therefore helps distinguish investigative financial analysis from formal expert examination, but leaves open questions concerning the procedural treatment of such reports, the parties' ability to examine and challenge them, and the basis on which courts should assess their evidentiary significance.

These questions will be addressed in the following Section. The assessment will therefore focus on whether the Serbian framework provides a coherent relationship between investigative financial analysis, formal expert examination and party associated professional assistance, and whether the procedural safeguards applicable to each are sufficient to ensure effective judicial scrutiny and procedural fairness.

3. Strengthening the Coherence of Forensic Accounting Expertise in Serbian Criminal Proceedings

The comparative analysis indicates that legal systems differ significantly in the institutional organisation of forensic accounting expertise, while the examined models all incorporate procedural safeguards concerning the production, presentation and evaluation of expert evidence. Party initiated expertise relies on disclosure, adversarial challenge and judicial assessment of reliability. Judicially controlled expertise places greater responsibility on the court for selecting and

structuring the examination. Investigative models integrate specialised financial knowledge into the investigation while requiring appropriate safeguards when that knowledge is subsequently relied upon as evidence. The relevant conclusion from the comparative analysis is therefore not that one particular institutional model should prevail, but that any model must clearly define the role of forensic accounting expertise and provide effective safeguards for assessing the reliability of expert evidence.

This finding is also relevant to Serbia. The Serbian legal system already contains the three principal modalities examined in this paper: formal expert examination under the provisions of the CPC, professional expert assistance selected by the parties, and specialised financial forensic expertise embedded within the public prosecution structure. The issue is therefore not simply the existence of these institutional modalities, but their practical capacity and the clarity of their procedural relationship. As discussed above, the professional consultant's effectiveness depends on meaningful participation in expert examination and the ability to contribute to the critical assessment of expert evidence, while the functioning of prosecutorial financial forensic expertise is affected by staffing and integration of financial forensic experts and by the capacity of prosecutors to identify, understand and effectively use specialised financial expertise. The respective functions of these modalities can be identified, but their interaction when different forms of financial expertise are present in the same proceedings is not fully articulated by the existing framework. To illustrate, in a financial crime case involving allegedly misappropriated funds, a financial forensic expert within the prosecution may analyse financial records during the investigation, while the defence may subsequently engage a professional consultant who raises objections to that analysis or assists in assessing formal expert findings. If the court later orders formal expert examination, the procedural framework must be applied in a manner that enables the court and the parties to distinguish the different functions of these forms of expertise and to assess their respective evidentiary significance.

The first priority area for reform is the legal regulation of formal expert examination, particularly the maintenance and assessment of professional competence after appointment. The Serbian Law on Court Experts provides the institutional basis for the appointment and registration of experts and establishes requirements relating to professional competence. These requirements are important, but initial qualification for appointment does not by itself ensure that an expert remains adequately qualified to perform complex forensic examinations throughout the period in which they are registered. This is particularly significant in financial crime proceedings, where the complexity and methods of fraud may evolve alongside changes in accounting standards, financial practices, investigative techniques and analytical methods. The distinction is therefore between eligibility for appointment and continuing competence to perform forensic work. A more robust framework should consequently provide for appropriate mechanisms for maintaining and periodically assessing professional competence, including, where appropriate, continuing

professional development and reassessment of competence, rather than treating qualification as a condition that is effectively established once and then presumed to continue.

Expert evidence creates challenges for criminal adjudication because courts must rely on specialised knowledge that they are not themselves equipped to verify independently. This tension is identified in the Serbian context by Soković (2008), who emphasises the relationship between the court's responsibility for determining the facts and the specialised knowledge possessed by the expert, which may substantially influence that determination. This problem is not unique to Serbia. Comparative sources similarly show that neither judicial appointment nor adversarial presentation provides an automatic guarantee of reliable expert evidence. The relevant safeguards must concern the quality of expertise and the procedural conditions under which it is tested (Decaigny, 2014; Roberts, 2015).

For financial forensic expertise, this concern is particularly concrete. The reliability of an expert opinion may depend on the completeness and relevance of the financial records examined, the reconstruction of transactions, the assumptions applied, the analytical methods employed, and the reasoning connecting the underlying evidence to the final conclusion. Research conducted in the United States by Abukhalaf, Aldridge and Smith (2021), examining judicial decisions concerning forensic accountant testimony, identified deficiencies in evidence, methodology and reasoning among the principal grounds for excluding such testimony. For the purposes of the present analysis, this finding illustrates the importance of methodological and evidentiary scrutiny of forensic accounting opinions, rather than providing direct evidence concerning Serbian practice. The implication for Serbian law is that regulation of court experts should not be confined to whether a person satisfies the conditions for appointment. It should also provide an adequate institutional basis for specialisation, continuing professional competence, methodological quality and professional accountability.

The objective of reform should not be to strengthen the formal authority of court experts, but to ensure the effective operation of the procedural and professional safeguards already provided by the CPC and the Law on Court Experts. The fact that a person is registered as a court expert should not, by itself, determine the evidential value of the expert opinion. The CPC already provides mechanisms for addressing deficiencies in expert evidence. Under Article 123, parties may inspect the findings and opinion and submit observations, while Article 124 requires the procedural authority to order repeated expert examination where the findings or opinion are unclear, incomplete, erroneous or contradictory, and to appoint another expert where such deficiencies cannot otherwise be remedied. The Law on Court Experts complements these safeguards. Article 18 requires the court or other procedural authority to monitor the expert's work and permits the submission of a reasoned proposal for the expert's dismissal, while Article 19 provides for dismissal where an expert performs the examination improperly, negligently or unprofessionally, expressly including incomplete, unclear, contradictory or inaccurate findings.

Article 27 further requires experts to perform their work conscientiously, professionally and impartially.

The relevant reform question is therefore how the existing mechanisms for addressing deficient expert work can be complemented by clearer requirements concerning the maintenance and assessment of professional competence over time. Particular attention may be given to the consistent use of Article 124 CPC when deficiencies are identified and to the practical operation of the monitoring and accountability mechanisms under Articles 18 and 19 of the Law on Court Experts. The same perspective supports greater attention to specialisation and continuing professional development where the complexity of financial crime requires expertise extending beyond general qualifications in accounting, economics or finance. This would shift the focus of reform from creating additional formal controls to ensuring that existing mechanisms provide an effective basis for assessing the quality and reliability of expert evidence in individual criminal proceedings.

A related issue concerns the practical effectiveness of the professional consultant. The Serbian CPC provides parties with a procedural mechanism for obtaining specialised assistance in connection with expert examination, but the formal availability of that mechanism does not necessarily ensure effective participation in practice. As discussed above, Serbian literature indicates that the professional consultant is primarily engaged in practice by the defence and may assist the accused in understanding expert findings, formulating objections and addressing specialised issues arising from the examination (Janković, 2025). The practical significance of the institute therefore depends on the consultant's ability to participate meaningfully in the examination and to contribute to the critical assessment of expert evidence. Accordingly, the effectiveness of this modality could be strengthened by ensuring that the procedural opportunities available to the professional consultant are capable of being exercised in practice and that observations concerning expert findings are given appropriate procedural consideration. This would reinforce the adversarial examination of expert evidence without transforming the professional consultant into a privately appointed court expert.

The second reform area concerns the evidentiary implications of financial forensic expertise within the public prosecution structure. The comparative analysis shows that integration of forensic accounting expertise into the investigative function is not an exception. The English model demonstrates how specialised financial expertise may be incorporated into multidisciplinary investigation and prosecution of complex economic crime. The Serbian Financial Forensic Service performs a comparable investigative function within a different institutional structure. Its existence should therefore be regarded as an important component of the Serbian legal and institutional system for the suppression of organised crime and corruption rather than as an alternative to formal expert examination.

The difficulty arises from the procedural consequences of that investigative function. The Financial Forensic Service produces specialised financial analysis for the

purposes of assisting the public prosecutor. Its institutional purpose is consequently different from that of an expert appointed to provide findings and an opinion within the procedural framework of formal expert examination. Nevertheless, the analysis produced during the investigation may subsequently become relevant to the adjudication of the criminal case. This makes the procedural treatment of the resulting report particularly important. In the case Kzz 621/25, the Supreme Court of Serbia held that a report prepared by a financial forensic expert could be used as evidence in criminal proceedings and that its evidentiary strength was a matter for judicial assessment. The Court also held that such a report represents evidence and cannot replace economic and financial expert examination that had also been ordered and conducted in the case.

The decision therefore provides an important clarification of the evidentiary position of financial forensic analysis. It confirms that the report of a financial forensic expert is not excluded from evidentiary use merely because it originates from the prosecutorial financial forensic structure, while at the same time preserving the distinct procedural role of formal expert examination. The remaining issue is consequently not whether such reports can have evidentiary relevance, but how their procedural treatment should be articulated when they are relied upon alongside formal expert evidence or other competing financial analyses. Questions may arise, for example, concerning the manner in which the methodology and underlying financial material are disclosed, how the author of the report participates in the proceedings, and how the parties and the court can effectively examine and challenge the conclusions made.

The suggested reform is therefore not aimed at transforming financial forensic experts into court experts, since this would undermine the distinctive investigative function for which the Financial Forensic Service was established. Rather, the framework should more precisely articulate the procedural function of financial forensic analysis and the conditions under which its results may be relied upon and tested in criminal proceedings. This should include appropriate requirements concerning the methods and financial material on which the analysis is based, disclosure of the reasoning leading to its conclusions, and procedural opportunities enabling the parties and the court to examine and challenge those conclusions.

The third reform area concerns the interaction between financial forensic analysis and formal expert examination. The Serbian legal framework establishes different institutional and procedural bases for these two forms of specialised financial analysis. Under Article 19 of the Law on the Organisation and Jurisdiction of State Authorities in the Suppression of Organised Crime, Terrorism and Corruption, financial forensic experts assist the competent public prosecutor in the analysis of money flows and financial transactions for the purpose of criminal prosecution. Formal expert examination, by contrast, is governed by the CPC and is ordered when specialised knowledge is required for the establishment or assessment of a fact in the proceedings. The distinction therefore follows primarily from the different procedural functions through which the specialised knowledge is generated and used.

For the Serbian framework, the principal issue is consequently not the existence of two competing forms of expertise, but the need for a sufficiently clear procedural relationship between them when they address overlapping financial questions. A financial forensic analysis may assist the prosecutor in reconstructing transactions and identifying relevant financial relationships during the investigation, while formal expert examination may subsequently be ordered to establish or assess facts requiring specialised knowledge within the procedural framework of the CPC. Where the analyses differ, the procedural framework should enable the court to identify the source, purpose and methodological basis of each analysis and to assess them according to their respective procedural functions and evidentiary significance. This approach would preserve the investigative value of financial forensic expertise while maintaining the distinct role of formal expert examination.

Taken together, these issues indicate that the principal challenge for Serbian law is not the absence of institutional modalities or formal safeguards, but the need to strengthen their practical effectiveness and to articulate more clearly the procedural relationship between them. Formal expert examination requires effective mechanisms for maintaining professional competence and responding to deficient expert work. The professional consultant requires meaningful procedural participation in the examination and assessment of expert evidence. Financial forensic expertise within the public prosecution structure requires clearer articulation of the procedural treatment of its analytical outputs. Finally, where different forms of financial expertise coexist in the same proceedings, the procedural framework should enable their respective functions, methodological bases and evidentiary significance to be clearly distinguished and judicially assessed. Such clarification would strengthen the coherence of the Serbian framework without requiring the replacement of its existing hybrid institutional structure.

Conclusion

The analysis shows that Serbia's hybrid structure can be retained, but its legislative and institutional framework should be further clarified and strengthened. The following measures are proposed:

1. *Law on Court Experts*. Amend the Law to distinguish initial eligibility for appointment from continuing professional competence, including appropriate requirements for continuing professional development and maintenance of specialised knowledge.
2. *Criminal Procedure Code*. Amend the provisions concerning the professional consultant to ensure meaningful participation in the examination of expert evidence and effective consideration of observations and objections concerning expert findings and opinion.
3. *Criminal Procedure Code*. Clarify how financial forensic reports are to be treated in relation to formal expert findings and opinions, including their respective

procedural functions and evidentiary significance where they concern the same facts. The regulation should enable the court and the parties to distinguish their respective source, purpose and procedural function.

4. *Law on the Organisation and Jurisdiction of State Authorities in the Suppression of Organised Crime, Terrorism and Corruption*. Clarify the institutional and procedural role of the Financial Forensic Service, including the preparation and use of its financial analyses and reports.

5. *Institutional capacity*. Strengthen the capacity of judges and prosecutors to identify, understand and critically assess specialised financial expertise in complex financial crime proceedings.

These measures would preserve the existing hybrid structure while providing clearer procedural rules, stronger professional safeguards and more effective scrutiny of forensic accounting expertise.

References

- Abukhalaf, R., Aldridge, M., & Smith, J. (2021). Methodological Deficiencies in the Expert Testimony of Forensic Accountants: A Qualitative Content Analysis of Judicial Statements Pertaining to Daubert Exclusions. *Journal of Accounting, Finance & Auditing Studies*, 7(2), 109-132. <https://doi.org/10.32602/jafas.2021.016>
- Criminal Procedure Code of the Republic of Serbia, Official Gazette of the Republic of Serbia*, No. 72/2011, as amended, including No. 62/2021.
- Dcaigny, T. (2014). Inquisitorial and Adversarial Expert Examinations in the Case Law of the European Court of Human Rights. *New Journal of European Criminal Law*, 5(2). 149-166. <https://doi.org/10.1177/203228441400500203>
- Hegazy, S., Sangster A., & Kotb, A. (2017). Mapping Forensic Accounting in the UK. *Journal of International Accounting, Auditing and Taxation*, 28, 43-56. <https://doi.org/10.1016/j.intaccaudtax.2016.12.004>
- Janković, D. (2025). Stručni savetnik: Krivičnoprocesni položaj i sudska praksa. In M. Kolaković-Bojović & I. Stevanović (Eds.), *Veštačenje u kaznenim postupcima* (pp. 107–123). Institut za kriminološka i sociološka istraživanja. <https://doi.org/10.47152/Palic2025.6>
- Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993). <https://supreme.justia.com/cases/federal/us/509/579>
- Kumho Tire Co. v. Carmichael, 526 U.S. 137 (1999). <https://supreme.justia.com/cases/federal/us/526/137>
- Kolaković-Bojović, M., & Matić Bošković, M. (2025). Building prosecutorial capacity for suppression of corruption in Serbia. In D. Malinić & S. Vučković Milutinović (Eds.), *Contemporary forensic accounting trends in combating financial*

fraud: Global and national perspectives (pp. 245–258). University of Belgrade, Faculty of Economics and Business, Publishing Centre. <https://cfr.ekof.bg.ac.rs/wp-content/uploads/2025/10/Conference-Proceedings-Second-Conference-on-Forensic-Accounting.pdf>

Law on Expert Witnesses, *Official Gazette of the Republic of Serbia*, No. 44/2010.

Law on Organization and Jurisdiction of Government Authorities in Suppression of Organized Crime, Terrorism and Corruption, *Official Gazette of the Republic of Serbia*, No. 94/2016, 87/2018, 10/2023.

Levi, M. (2003). The Roskill Fraud Commission revisited: an assessment, *Journal of Financial Crime*, 11(1), 38–44. <https://doi.org/10.1108/13590790410809022>

Milojević, M. (2020). Usluge stručnog savetnika u krivičnom postupku. In: *Usluge i prava korisnika* (pp. 417–432). Pravni fakultet Univerziteta u Kragujevcu, Institut za pravne i društvene nauke. <https://doi.org/10.46793/UPK20.417M>

Roberts, P. (2015). Paradigms of forensic science and legal process: a critical diagnosis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1674), 20140256. <https://doi.org/10.1098/rstb.2014.0256>

Singleton, T. W., & Singleton, A. J. (2010). *Fraud auditing and forensic accounting* (4th ed.). John Wiley & Sons. <https://doi.org/10.1002/9781118269183>

Sinanović B., Marić B., Božović V., Stojadinović D., Nikolic G., Jakovljević J., Kostić J. (2021). Praktikum za istrage i suđenja krivičnih dela korupcije. USAID Projekat za odgovornu vlast. <https://www.pars.rs/public/Dokumenti/Publikacije/1151/Praktikum-za-istrage-i-sudjenja-KD-korupcije.pdf>

Soković S. (2008). Veštačenje u krivičnom postupku: nova praksa u starim normativnim okvirima i drugi problemi, *Zbornik Instituta za kriminološka i sociološka istraživanja* 27(1–2), 29–54. <https://zbornik-iksi.rs/index.php/home/article/view/99/181>

Code of Criminal Procedure (Strafprozeßordnung [StPO]). <https://www.gesetze-im-internet.de/stpo/BJNR006290950.html>

Criminal Justice Act 2003, c. 44. <https://www.legislation.gov.uk/ukpga/2003/44/contents>

Criminal Procedure Rules 2025, Statutory Instrument 2025 No. 909, Part 19 (Expert Evidence). <https://www.legislation.gov.uk/uksi/2025/909/part/19>

Federal Rules of Evidence, Rules 702, 703, and 705. <https://www.uscourts.gov/rules-policies/current-rules-practice-procedure>

Federal Rules of Criminal Procedure, Rules 16(a)(1)(G) and 16(b)(1)(C). <https://www.uscourts.gov/rules-policies/current-rules-practice-procedure>

Kzz 621/25 (Supreme Court of Serbia, November 25, 2025).
<https://www.vrh.sud.rs/sr-lat/sudska-praksa>

R v. H [2007] UKHL 7. <https://publications.parliament.uk/pa/ld200607/ldjudgmt/jd070228/rvh-1.htm>

Weigant G. (1976). Rechtsfindung durch Sachverständige? Zur Problematik von Gerichtsgutachten. *Aus Politik und Zeitgeschichte*, 44, 35-47.
<https://www.bpb.de/shop/zeitschriften/apuz/archiv/530296/rechtsfindung-durch-sachverstaendige-zur-problematik-von-gerichtsgutachten/>

Biljana Simić
S financial & tax consulting,
Belgrade
ORCID: 0009-0002-9181-8934

FROM DOCUMENTATION TO ECONOMIC SUBSTANCE: FORENSIC INDICATORS OF TAX RISKS AND TRANSFER PRICING MANIPULATION IN THE ERA OF ARTIFICIAL INTELLIGENCE

UDK 336.2:657.6

DOI <https://doi.org/10.54318/czf.2026.bs.028>

Abstract

The formal accuracy and internal consistency of contracts, invoices, calculations, and information contained in transfer pricing studies do not, in themselves, provide sufficient evidence that transactions between related parties faithfully reflect their economic substance, underlying business rationale, and prevailing market conditions. The purpose of this paper is to identify and systematise forensic indicators of tax risks and potential transfer pricing manipulation, while examining the contribution of AI tools to documentary and financial data analysis, anomaly detection, and the early identification of controlled transactions presenting an elevated level of risk.

The research combines an analysis of the relevant regulatory and forensic framework with a comparative review of anonymised cases from professional practice. Documentary, functional, and financial analyses are used to assess whether contractual arrangements correspond to the actual implementation of controlled transactions, including the functions performed, assets used, risks assumed and controlled, financial flows, economic benefits, and business rationale. Particular attention is given to unusual price and profitability deviations, inappropriate cost allocation, selective use of comparable data, service fees unsupported by appropriate evidence, and profit-shifting patterns.

The findings indicate that an isolated formal irregularity is generally insufficient for a reliable risk assessment. Its forensic relevance arises from the combined evaluation of documentary inconsistencies, financial anomalies, and discrepancies identified through functional analysis. Based on the findings, a structured framework is proposed for classifying indicators according to their source, significance, and interrelationship. Artificial intelligence tools can enhance the processing, integration, and analysis of documentary and financial data, as well as the identification of anomalies and elevated-risk patterns. However, their outputs constitute analytical input rather than a final forensic conclusion. Assessing the economic substance of a transaction, the materiality of identified deviations, and

their evidentiary relevance requires professional scepticism, critical evaluation of data sources, and independent validation by a qualified financial forensic expert.

Key words: *financial forensics, tax risks, transfer pricing, economic substance, functional analysis, AI tools, forensic indicators*

JEL Classification: *M42, H26, C55*

Introduction

Transfer pricing represents one of the most complex areas of taxation involving transactions between associated enterprises, as the conditions under which controlled transactions are conducted affect the allocation of profits between associated enterprises and, consequently, the determination of the taxable base. Transfer pricing is, in itself, a legitimate feature of transactions between associated enterprises; however, where the conditions of controlled transactions differ from those that would have been agreed between independent enterprises in comparable circumstances, taxable profits may be inappropriately allocated, giving rise to increased tax risk. The arm's length principle therefore constitutes the internationally accepted standard for evaluating transactions between associated enterprises, as reflected in both the OECD Transfer Pricing Guidelines and the United Nations Practical Manual on Transfer Pricing (OECD, 2022; United Nations, 2021). Contemporary transfer pricing analysis, however, is not confined to formally agreed contractual terms. It requires consideration of the economically relevant characteristics of the transaction, including the functions actually performed by the parties, the assets used, the risks assumed and controlled, and the relevant economic circumstances. It is on this basis that the OECD Guidelines develop the concept of the accurate delineation of the actual transaction, through which formally documented arrangements are assessed in conjunction with the actual conduct of the associated enterprises and the economic circumstances surrounding the transaction (OECD, 2022).

Transfer pricing compliance, however, cannot be reduced to the existence of formally accurate contracts, invoices, calculations, and information contained in transfer pricing documentation. A reliable analysis requires an assessment of whether the documented arrangements correspond to the actual implementation of the controlled transaction, including the relevant financial flows, the allocation of economic benefits, and the underlying business rationale. Transfer pricing documentation therefore constitutes an important source of information and evidentiary support for the analysis, but its formal completeness and internal consistency are not, in themselves, sufficient to establish that the documented arrangement faithfully reflects the actual economic substance of the controlled transaction.

A discrepancy between the formally documented arrangement and the actual economic substance of a transaction represents a particularly relevant area for

forensic analysis. The forensic approach does not presume that every deviation constitutes tax fraud or manipulation; rather, it seeks to identify and assess documentary, functional, and financial indicators that may signal an elevated level of tax risk. Particular relevance may arise from inconsistencies between the contractual allocation of functions and risks and their actual performance and control, unusual movements in prices and profitability, inappropriate cost allocation, selective use of comparable data, service fees unsupported by appropriate evidence of the economic benefit, and other indicators suggesting a potential misalignment between the allocation of taxable profits and the economic substance of the controlled transaction. An individual anomaly, however, does not itself provide a sufficient basis for a forensic conclusion; its relevance depends on materiality, the quality and reliability of the supporting evidence, and its relationship with other available information (Trotman & Wright, 2012; Nelson, 2009).

The increasing digitalisation of business processes further expands the analytical possibilities in this area. Artificial intelligence tools can support the processing and integration of large volumes of documentary and financial data, the identification of unusual patterns, and the early detection of transactions presenting an elevated level of risk. Their role, however, remains analytical and supportive: assessing the actual economic substance of a transaction, the materiality of identified deviations, and their evidentiary relevance requires professional scepticism, critical evaluation of data sources, and independent validation of the findings (OECD, 2025; Nelson, 2009).

Against this background, the purpose of this paper is to identify and systematise forensic indicators of tax risks and potential transfer pricing manipulation by linking formal documentation with the actual economic substance of controlled transactions. The research combines an analysis of the relevant regulatory and forensic framework with a comparative examination of selected anonymised cases drawn from professional practice, analysed through a documentary, functional, and financial framework.

The contribution of the paper lies in proposing a structured framework for classifying forensic indicators according to their source, significance, and interrelationship. The framework is based on the premise that the convergence of multiple independent and materially relevant indicators pointing to the same economic inconsistency increases the forensic relevance of the findings and provides a more robust basis for further tax and forensic examination. The framework developed in this study draws on the regulatory concept of the economically relevant characteristics of controlled transactions, professional scepticism, and evidence triangulation, while its application is examined through selected anonymised cases from professional practice.

The study builds on related-party transaction literature (Ryngaert & Thomas, 2012) and forensic transfer pricing research (Curtis, 2016; Curtis & Lahav, 2020), integrating them through a structured evidentiary assessment.

1. Economic Substance of Controlled Transactions

The arm's length principle requires controlled transactions to be analysed according to their economically relevant characteristics rather than solely on the basis of supporting documentation. The accurate delineation of the actual transaction therefore requires consideration of the contractual terms, functions actually performed, assets used, economically significant risks assumed and controlled, relevant economic circumstances, and business strategies pursued by the parties. In relation to risk, contractual allocation must be assessed against actual conduct, including which enterprise exercises control over the risk and has the financial capacity to assume it (OECD, 2022; United Nations, 2021).

Transfer pricing documentation may therefore be formally complete while remaining analytically insufficient where the functional profile, risk allocation, selection of comparables, or remuneration methodology is inconsistent with the underlying facts. For the purposes of this study, documentary compliance refers to the completeness and internal consistency of information supporting a transfer pricing position, whereas substantive alignment requires the reported outcome to be consistent with the accurately delineated transaction and its economically relevant characteristics. Documentation is therefore treated as an important evidentiary source rather than conclusive evidence that the reported outcome satisfies the arm's length principle (OECD, 2022).

This distinction is also reflected in academic research. Transfer pricing should not be viewed solely as a technical mechanism for allocating costs and revenues within a group, since it may materially affect the allocation of taxable profits (Sikka & Willmott, 2010). Empirical research further indicates that transfer pricing documentation requirements may influence profit-shifting behaviour, although their effectiveness varies according to firm-specific characteristics and the complexity of multinational structures (Beer & Loeprick, 2015). These findings reinforce the need to assess the economic effects of controlled transactions together with their formal documentation.

Academic research also shows that related-party transactions are heterogeneous. Ryngaert and Thomas (2012) distinguish *ex ante* RPTs, which predate related-party status, from *ex post* RPTs initiated afterwards. They report different economic associations, with *ex post* RPTs linked to lower operating profitability and adverse outcomes. This supports context-sensitive assessment rather than treating related-party status as determinative of risk.

For forensic transfer pricing assessment, related-party status is contextual evidence rather than a conclusion. Analysis should focus on actual conduct, functions and risks, timing, and consistency between documentation and economic outcome.

2. Forensic Assessment of Transfer Pricing Risk: From Isolated Anomalies to the Convergence of Indicators

2.1. From Anomaly to Forensic Relevance

Forensic analysis distinguishes between anomalies, risk indicators, and forensic conclusions. An anomaly is a departure from an expected economic, financial, transactional, or documentary pattern; it becomes a relevant risk indicator only when its materiality and context warrant further examination. A forensic conclusion requires a higher degree of evidentiary support and cannot rest on a single red flag. Professional scepticism therefore requires consideration of whether a discrepancy reflects legitimate commercial circumstances, error, inadequate documentation, or potential intentional misstatement (International Auditing and Assurance Standards Board [IAASB], 2025; Kaur et al., 2023).

This distinction is particularly important in transfer pricing, where unusual prices, margins, contractual terms, or cost allocations may have more than one economically plausible explanation. The forensic relevance of an anomaly therefore depends on its materiality, economic consequences, supporting evidence, and consistency with the broader circumstances of the controlled transaction.

2.2. Evidential Corroboration and the Convergence of Indicators

The significance of a risk indicator depends on whether it is corroborated, contradicted, or remains unsupported by other evidence. Research on evidence triangulation shows that fraud-risk assessments are affected by the consistency and relative independence of available evidence and that contradictory evidence may increase professional scepticism toward management-controlled information (Trotman & Wright, 2012). Transfer pricing discrepancies should therefore be assessed against the broader documentary, functional, financial, and transactional record.

For the purposes of this study, indicator convergence means that analytically distinct findings from sufficiently independent sources point to the same underlying economic inconsistency or tax consequence. Its relevance depends on evidentiary reliability and independence, materiality, potential financial and tax effects, and plausible alternative commercial explanations. Convergence therefore requires evidence-based assessment rather than mechanical accumulation or scoring of red flags.

2.3. A Structured Framework for Forensic Risk Assessment: Source, Significance and Interrelationship

The source–significance–interrelationship framework is developed in this study as a qualitative approach to evaluating discrepancies identified in controlled transactions. It does not replace transfer pricing analysis or constitute a general model of

fraudulent behaviour; its purpose is to structure the evidentiary assessment of indicators once potential inconsistencies have been identified.

The framework is informed by three complementary strands. First, transfer pricing analysis requires consideration of economically relevant characteristics and actual conduct rather than contractual form alone (OECD, 2022). Second, professional scepticism requires critical evaluation of the reliability, sufficiency, and consistency of evidence (Nelson, 2009). Third, evidence-triangulation research emphasises corroboration across sufficiently independent sources rather than reliance on isolated indicators (Trotman & Wright, 2012). Broader fraud-risk models seek primarily to explain conditions or behavioural factors associated with fraudulent conduct; the framework proposed here instead structures transaction-specific evidentiary assessment and does not infer motivation or intent (Dorminey et al., 2012).

Forensic transfer pricing also has an academic basis. Curtis (2016) links forensic accounting, economics, and data analytics to transfer pricing compliance and enforcement, while Curtis and Lahav (2020) illustrate such approaches through case studies.

The framework builds on that literature but serves a different purpose: it structures transaction-level discrepancies by source, significance, and interrelationship rather than inferring manipulation. Relevance depends on evidence quality, materiality, convergence, and plausible alternative explanations.

Each indicator is assessed across three dimensions. Source considers the origin, reliability, and relative independence of documentary, functional, financial, transactional, and relevant contextual information. Significance considers whether a discrepancy may materially affect the economic interpretation of the transaction, remuneration, taxable-profit allocation, or the arm's length outcome. Interrelationship considers whether separate findings corroborate, contradict, or remain neutral toward the same economic explanation. Forensic relevance is strengthened where material findings from sufficiently independent sources converge and cannot be adequately explained by legitimate commercial circumstances, while contradictory evidence and alternative explanations must also be considered.

The framework therefore produces a structured risk profile rather than a binary conclusion on manipulation. It supports comparison across the selected cases and identifies controlled transactions warranting enhanced tax or forensic examination.

3. Research methodology

The study employs a qualitative comparative multiple-case design to examine how forensic indicators of transfer pricing risk emerge and interact across selected controlled transactions from professional practice. This approach is appropriate because assessment of economic substance and evidentiary consistency is fact-dependent and requires multiple sources of evidence within their business context.

The cases are used for analytical comparison rather than statistical generalisation (Eisenhardt, 1989; Yin, 2018).

Cases were selected purposively using two complementary criteria: availability of sufficiently robust documentary and financial evidence and the presence of potentially material discrepancies between the documented transfer pricing position and the transaction as implemented. Case I concerns medium-sized wholesale and retail enterprises and a business and technical cooperation arrangement involving inventory, warehousing, financing, product placement, development activities, and related commercial relationships. Case II concerns small enterprises subject to statutory audit due to turnover, operating within a predominantly family-related network involving IT services, commercial real estate leasing, software development, and interconnected ownership, management, and investment relationships.

The evidentiary record includes transfer pricing studies, contracts, invoices, accounting records, bank statements, financial information, and other transaction-specific documentation. Multiple interviews were conducted in both cases as supplementary evidence. In Case I, interviewees included the director, owner, CFO, representatives of finance, procurement and sales, legal advisers, and tax consultants, selected according to their functional involvement, business-process responsibilities, and role in approving contractual arrangements. Topics included business rationale, allocation of functions and risks, customer relationships and deposits, insurance arrangements, and development funding. In Case II, interviews with the CFO, management representatives, and owner addressed business activities, commercial rationale, and the nature of related-party relationships. Interviews were documented through contemporaneous notes and checked against documentary and financial evidence.

Interview evidence was treated as contextual and supplementary rather than independently determinative. Greater evidentiary weight was given to documentary and financial records where statements could not be corroborated; behavioural or contextual information was not used independently to infer intent, manipulation, or tax evasion. Identifying and commercially sensitive information was excluded or modified while preserving facts material to the analysis.

The analytical procedure combines within-case examination and cross-case comparison. Each transaction is reconstructed to identify material documentary, functional, financial, and transactional discrepancies, which are then assessed by source, significance, and interrelationship, including evidentiary reliability, potential effects on remuneration and taxable-profit allocation, corroboration, and plausible alternative explanations. Cross-case comparison examines whether recurring combinations of indicators converge on similar inconsistencies.

The purposive, anonymised, two-case design and confidentiality constraints limit statistical generalisation. The findings therefore support analytical interpretation

rather than estimates of prevalence, and these limitations are taken into account when assessing the scope of the results.

4. Comparative Forensic Analysis of Selected Controlled Transactions

The analysis now applies the proposed framework to selected anonymised cases from professional practice, examining formally documented arrangements against the parties' actual conduct and observed economic and financial consequences.

4.1. Case I: Revenue-Sharing, Functional Contributions and the Allocation of Economic Returns

Case I concerns a business and technical cooperation agreement between two related companies. Entity A contributed existing inventory, warehousing facilities, and related infrastructure, while Entity B, established only a few days before the agreement, was contractually described as contributing marketing knowledge and methodology, financing acquisitions, product placement, and debt collection. Entity B had pricing authority subject to a cap based on prices previously achieved by Entity A, while the agreement provided for a 50:50 revenue split. This contractual symmetry was examined against the parties' actual economic contributions.

Actual conduct indicated a more differentiated allocation of functions and exposures. Entity A retained ownership of the goods and bore contractual risk of loss or deterioration, while Entity B financed acquisitions, employed personnel, and conducted commercial activities. Entity A had also received customer deposits before Entity B entered the arrangement, so relevant customer relationships and commercial commitments predated Entity B's formal participation.

The analysis extended beyond the principal agreement. A separate development agreement required Entity A to develop an innovative product for Entity B, raising the question whether the functions, resources, costs, outputs, and benefits supported the agreed remuneration. This is consistent with academic work on intra-group services. Plesner Rossing and Pearson (2022) emphasise arm's-length pricing and relevant cost-allocation considerations. Here, this supports testing whether functions performed, resources used, and costs incurred are consistent with the agreed remuneration. Discrepancies were also identified concerning an alleged insurance arrangement with Entity C. These relationships were considered both separately and for their combined effect on costs and economic returns.

Interview evidence, checked against available documentation, provided a plausible commercial context for part of Entity A's development activity: Entity A was a beneficiary of foreign development funding related to innovative products in raw wheat processing. This circumstance neither established nor excluded manipulation. Taken together, customer deposits, inventory ownership, infrastructure, differentiated exposures, development arrangements, and insurance-related discrepancies indicate a potential asymmetry between revenue allocation and

economic contributions. They do not establish manipulation or tax evasion, but support further examination of whether Entity B's remuneration is commensurate with its contribution and whether costs recognised by Entity A correspond to economically useful activities.

4.2. Case II: Preferential Lease Terms, Related-Party Networks and the Allocation of Economic Benefits

Case II concerns a related-party network developed after the reorganisation of Entity A, which combined IT services with commercial real estate leasing. Entity B was created through a spin-off by incorporation, initially owned 30% by Entity A and 70% by Entity A's founder, and later wholly owned by that individual. Entity B subsequently leased premises to Entity C, owned by the founder's son, at a price significantly lower than that charged to unrelated tenants. The lease was indefinite, required no deposit, and placed investment and technical maintenance costs on the lessor. Because Entity B had comparable leases with unrelated tenants, these transactions provide a relevant starting point for an internal comparable uncontrolled price analysis, subject to adjustment for material contractual and economic differences (Republic of Serbia, Ministry of Finance, 2021).

Management roles also overlapped: the founder of Entity C was an authorised representative of Entity B, while the founder of Entity B was a director of Entity C. These relationships do not in themselves establish a departure from the arm's length principle, but, together with ownership and family relationships, are relevant to whether the lease terms reflect independent commercial negotiation (Republic of Serbia, Ministry of Finance, 2024).

Broader empirical literature links related-party pricing to ownership and governance. Almutairi et al. (2023), using Australian listed firms, report an association between tax-haven use and related-party sales pricing, alongside relationships involving earnings management and ownership concentration. Although tax havens are not directly relevant here, the study supports assessing pricing within its wider ownership and governance context. Accordingly, the lease-price difference is assessed within the ownership context, without treating it alone as evidence of an arm's-length departure.

The broader structure included software development transactions in which Entity C invoiced Entity A, which then re-invoiced end customers, while available documentation did not clearly allocate ownership, licensing, or rights to use the software. This raises questions as to which entity performs and controls the economically significant development functions, bears relevant risks, and is entitled to corresponding returns. Entity B also held a 50% interest in a real estate development project with the spouse of Entity C's founder under a joint development agreement.

Case II therefore shows that transfer pricing risk may arise through the cumulative effect of price differences, preferential contractual terms, management

interdependence, unclear allocation of rights, and interconnected investments. None of these findings independently establishes manipulation; their convergence supports examination of whether economic benefits are allocated at arm's length and whether the documentation reflects the economic substance of the relationships.

4.3. Cross-Case Analysis: Convergence of Forensic Indicators

The two cases show that elevated transfer pricing risk may emerge through different transactional configurations while the forensic logic remains consistent. In Case I, the principal discrepancy concerns the relationship between contractual revenue allocation and the distribution of functions, risks, assets, costs, and pre-existing commercial positions. In Case II, the economic effect is more dispersed across preferential terms, internal pricing differences, overlapping governance roles, unclear software-related rights, and interconnected property arrangements.

Across both cases, documentary evidence defines the formal structure, while functional, financial, transactional, and contextual evidence tests whether that structure corresponds to the parties' actual economic position. An indicator gains significance where a discrepancy may materially affect remuneration, costs, taxable profit, or economic benefits, and its evidentiary relevance is strengthened where sufficiently independent sources converge on the same inconsistency. Plausible commercial explanations may reduce or qualify the significance of individual indicators and must therefore form part of the assessment.

The cross-case findings support a risk-based rather than binary interpretation. Below-market prices, unusual contractual terms, ownership changes, unsupported costs, or functional discrepancies are not sufficient individually to establish manipulation. Forensic relevance arises where material findings are supported by reliable evidence and interrelate in a coherent economic explanation. The comparative analysis thus operationalises the source–significance–interrelationship framework and shows why converging indicators provide a stronger basis for enhanced tax or forensic examination than isolated formal irregularities.

5. Artificial Intelligence as Analytical Support in Forensic Transfer Pricing Assessment

The relevance of artificial intelligence for forensic transfer pricing analysis lies primarily in its ability to integrate fragmented evidence rather than merely detect isolated anomalies. Controlled transactions generate information across contracts, invoices, accounting records, payment flows, transfer pricing documentation, ownership structures, and operational data. AI-supported tools can assist in extracting, classifying, reconciling, and time-stamping such information, facilitating transaction reconstruction and the identification of relationships that may remain less visible in sequential manual review. The increasing use of AI by tax administrations for analytical activities, compliance management, and risk-based case selection confirms the practical relevance of these capabilities (OECD, 2025).

For this study, the principal contribution of AI is evidence integration. Natural language processing can assist in identifying repetitive or inconsistent contractual clauses; financial analytics can highlight unusual prices, margins, costs, or payment patterns; and network analysis can map relationships among enterprises, individuals, contracts, and financial flows. AI can therefore support the source–significance–interrelationship framework by linking indicators from different sources and assessing whether they converge on the same underlying economic inconsistency.

The cases illustrate this potential. In Case I, AI-supported reconstruction could integrate customer deposits, the timing of Entity B’s establishment, inventory ownership, revenue sharing, risk allocation, development activities, insurance-related costs, and financial flows. In Case II, it could combine internal comparable rents, contractual terms, ownership and management relationships, software development transactions, and real estate investments.

AI-generated findings nevertheless remain analytical inputs rather than forensic conclusions. Correlations, price discrepancies, contractual similarities, or network relationships do not establish economic substance, intent, or manipulation. Their evidentiary value depends on data quality, transparency, contextual interpretation, corroboration, and professional judgement. AI can extend analytical scope, but materiality, arm’s length consistency, and evidentiary relevance remain subject to professional scepticism, qualified forensic assessment, and independent validation (IAASB, 2025).

Conclusion

This study shows that formal transfer pricing documentation, while necessary, cannot by itself establish that a controlled transaction reflects its actual economic substance. The cases indicate that elevated risk may arise through revenue-allocation asymmetries, preferential contractual terms, management interdependence, cost recognition, unclear allocation of rights, and fragmented arrangements. The forensic significance of such findings depends on materiality, reliability, and convergence across sufficiently independent sources rather than isolated anomalies.

The paper’s main contribution is the source–significance–interrelationship framework, which structures the assessment of forensic indicators without mechanical scoring. This positions the paper within related-party and forensic transfer pricing literature while distinguishing its contribution: structured evidentiary assessment of documentary, functional, financial, and transactional indicators. It supports examination of transfer pricing risk beyond nominal prices by considering contractual arrangements, functions, risks, rights, costs, timing, and related-party relationships. AI may strengthen evidence integration, transaction reconstruction, pattern identification, and risk prioritisation, but does not replace professional scepticism, contextual interpretation, or independent forensic judgement.

The purposive, anonymised two-case design and confidentiality constraints limit statistical generalisation; the contribution is analytical. The framework has therefore been operationalised through comparative case analysis rather than empirically validated in a statistical sense. Further research could examine inter-practitioner consistency and whether AI-supported evidence integration improves identification of material risk patterns without increasing false-positive assessments.

References

- Almutairi, A., Eulaiwi, B., Evans, R., & Taylor, G. (2023). Tax haven use and related-party transactions: Evidence from Australia. *Australian Accounting Review*, 33(4), 352–374. <https://doi.org/10.1111/auar.12413>
- Beer, S., & Loeprick, J. (2015). Profit shifting: Drivers of transfer (mis)pricing and the potential of countermeasures. *International Tax and Public Finance*, 22(3), 426–451. <https://doi.org/10.1007/s10797-014-9323-2>
- Curtis, S. L. (2016). Forensic approaches to transfer pricing compliance and enforcement. *Journal of Forensic & Investigative Accounting*, 8(3), 359–400. <https://ssrn.com/abstract=2786779>
- Curtis, S. L., & Lahav, Y. (2020). Forensic approaches to transfer pricing enforcement could restore billions in lost U.S. federal and state tax losses: A case study approach. *Journal of Forensic & Investigative Accounting*, 12(2), 284–315. <https://ssrn.com/abstract=3528318>
- Dorminey, J., Fleming, A. S., Kranacher, M., & Riley, R. A. (2012). The evolution of fraud theory. *Issues in Accounting Education*, 27(2), 555–579. <https://doi.org/10.2308/iace-50131>
- Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, 14(4), 532–550. <https://doi.org/10.5465/amr.1989.4308385>
- International Auditing and Assurance Standards Board. (2025). *How the IAASB's revised going concern and fraud standards reinforce professional skepticism*. International Auditing and Assurance Standards Board.
- Kaur, B., Sood, K., & Grima, S. (2023). A systematic review on forensic accounting and its contribution towards fraud detection and prevention. *Journal of Financial Regulation and Compliance*, 31(1), 60–95. <https://doi.org/10.1108/JFRC-02-2022-0015>
- Nelson, M. W. (2009). A model and literature review of professional skepticism in auditing. *Auditing: A Journal of Practice & Theory*, 28(2), 1–34. <https://doi.org/10.2308/aud.2009.28.2.1>
- OECD. (2022). *OECD transfer pricing guidelines for multinational enterprises and tax administrations 2022*. OECD Publishing. <https://doi.org/10.1787/0e655865-en>

OECD. (2025). *Tax administration 2025: Comparative information on OECD and other advanced and emerging economies*. OECD Publishing. <https://doi.org/10.1787/cc015ce8-en>

Plesner Rossing, J. C., & Pearson, T. C. (2022). Tax-compliant transfer pricing of intra-group services: The soft drink case. *Journal of Accounting Education*, 61, 100815. <https://doi.org/10.1016/j.jaccedu.2022.100815>

Republic of Serbia, Ministry of Finance. (2021). Rulebook on transfer pricing and methods applied under the arm's length principle in determining prices of transactions between related parties [Pravilnik o transfernim cenama i metodama koje se po principu „van dohvata ruke” primenjuju kod utvrđivanja cene transakcija među povezanim licima] (“Official Gazette of the Republic of Serbia”, Nos. 61/2013, 8/2014, 94/2019, and 95/2021).

Republic of Serbia, Ministry of Finance. (2024). Corporate Income Tax Law [Zakon o porezu na dobit pravnih lica] (“Official Gazette of the Republic of Serbia”, Nos. 25/2001, 80/2002, 43/2003, 84/2004, 18/2010, 101/2011, 119/2012, 47/2013, 108/2013, 68/2014, 142/2014, 91/2015, 112/2015, 113/2017, 95/2018, 86/2019, 153/2020, 118/2021, and 94/2024).

Ryngaert, M., & Thomas, S. (2012). Not all related party transactions (RPTs) are the same: Ex ante versus ex post RPTs. *Journal of Accounting Research*, 50(3), 845–882. <https://doi.org/10.1111/j.1475-679X.2012.00437.x>

Sikka, P., & Willmott, H. (2010). The dark side of transfer pricing: Its role in tax avoidance and wealth retentiveness. *Critical Perspectives on Accounting*, 21(4), 342–356. <https://doi.org/10.1016/j.cpa.2010.02.004>

Trotman, K. T., & Wright, W. F. (2012). Triangulation of audit evidence in fraud risk assessments. *Accounting, Organizations and Society*, 37(1), 41–53. <https://doi.org/10.1016/j.aos.2011.11.003>

United Nations, Department of Economic and Social Affairs. (2021). *United Nations practical manual on transfer pricing for developing countries 2021*. United Nations.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.

CIP - Каталогизација у публикацији
Народна библиотека Србије, Београд

657.632(082)
343.983:343.53(082)
502.131.1:658(082)

**CONFERENCE on forensic accounting From accounting
manipulation to sustainability and artificial intelligence: new horizons
in forensic accounting (3 ; 2026 ; Beograd)**

Conference proceedings / Third conference on forensic accounting From
accounting manipulation to sustainability and artificial intelligence: new
horizons in forensic accounting, Belgrade, October 2026. ; editors Dejan
Malinić, Savka Vučković Milutinović. - Beograd : University of Belgrade,
Faculty of economics and business, Publishing centre, 2026 (Beograd :
Službeni glasnik). - XII, 525 str. : graf. prikazi, tabele ; 24 cm

Tiraž 150. - Str. IX-XII: Foreword / editors. - Bibliografija uz svaki rad.

ISBN 978-86-403-1935-5

1. Malinić, Dejan, 1962- [уредник] [аутор додатног текста]
а) Форензичко рачуноводство -- Зборници б) Привредни криминал
-- Превенција -- Зборници в) Одрживи развој -- Економски аспект --
Зборници

COBISS.SR-ID 203249673