



Ferdowsi University of Mashhad
Press (FUMP)



IRANIAN JOURNAL OF ACCOUNTING, AUDITING and FINANCE

Volume 10, Issue 3, Summer 2026

IJA AF

Serial Number: 36

Print - ISSN: 2717-4131
Online - ISSN: 2588-6142

Title	Authors	Page Number
Sectoral Herding and Volatility in the Chinese Stock Market: Lessons from COVID-19	Syed Faisal Hassan Bukhari Pedro Pimentel Gualter Couto	1-19
Ranking the Factors Affecting the Level of Tax Compliance within the Framework of Comprehensive Tax Plan	Rahman Sajadpour Mehdi Arab Salehi Iraj Kazemi	21-45
Greener Growth, Tighter Credit? The Dual Role of Climate Action in Revenue Expectations and Financing Obstacles	Ly Hoang Vu	47-77
Macroprudential Policy and the Risk Exposure of Commercial Banks	Nafiseh Keshtgar Seyed Hossein Mirjalili	79-93
The Impact of ESG Performance on Capital Market Sensitivity	Behzad Kardan Hamideh Nazaridavaji	95-111
The Effectiveness of Kahoot! Game in Teaching Accounting Principles Course on Students' Quantitative and Qualitative Academic Performance	Mohammad Mahdi Bozorg Zohreh Arefmanesh Ahmad Zandvanian	113-137
The Moderating Effect of Power Distance on the Relationship between Conscientiousness and Extraversion with Auditors' Impartiality	Hamid Alviri Chenari Omid Pourheidari Ahmad Khodamipour	139-161
Board Characteristics, Audit Fees, and Political Connections: An Integration of Resource Dependence and Agency Theories	Ali Sayadi Iraj Mirhossini Reza Khajezadeh Saman Mirabbasi	163-185
The Relationship Between Managerial Overconfidence and ESG: Emphasizing The Role of Environmental Costing	Hanan Abdullah Hasan	187-200
Financial Stress Research Under the Magnitude Mind Lens	Taqi Al Abdwani Shima Ebrahimi Iliya Pishghadam	201-221

<https://ijaaf.um.ac.ir>

In

the Name of God, the Compassionate, the Merciful



Ferdowsi University of Mashhad

Iranian Journal of Accounting, Auditing & Finance (IJAAF)

Founder:

Ferdowsi University of Mashhad

Director:

Mohammad Ali Bagherpour Velashani, PhD

Editor in-Chief:

Mahdi Salehi, PhD

Journal Manager:

Solmaz Arefiasl

Monir Taheri

Frequency

Quarterly

Volume 10, Number 3, Summer 2026, Serial 36

ISSN (Print): 2717-4131, ISSN (Online): 2588-6142

Editorial Office: Faculty of Economics and Administrative sciences, Ferdowsi University of Mashhad, Azadi Sq., Mashhad; IRAN; P.O. Box: 1793; Postal Code: 9177948974; Tel: +98 51 38803742; Fax: +98 51 38763852

<http://ijaaf.um.ac.ir>

E-mail: ijaaf@um.ac.ir

Table of Contents

Title	Authors	Page
Sectoral Herding and Volatility in the Chinese Stock Market: Lessons from COVID-19	Syed Faisal Hassan Bukhari Pedro Pimentel Gualter Couto	1
Ranking the Factors Affecting the Level of Tax Compliance within the Framework of Comprehensive Tax Plan	Rahman Sajadpour Mehdi Arab Salehi Iraj Kazemi	21
Greener Growth, Tighter Credit? The Dual Role of Climate Action in Revenue Expectations and Financing Obstacles	Ly Hoang Vu	47
Macroprudential Policy and the Risk Exposure of Commercial Banks	Nafiseh Keshtgar Seyed Hossein Mirjalili	79
The Impact of ESG Performance on Capital Market Sensitivity	Behzad Kardan Hamideh Nazaridavaji	95
The Effectiveness of Kahoot! Game in Teaching Accounting Principles Course on Students' Quantitative and Qualitative Academic Performance	Mohammad Mahdi Bozorg Zohreh Arefmanesh Ahmad Zandvanian	113
The Moderating Effect of Power Distance on the Relationship between Conscientiousness and Extraversion with Auditors' Impartiality	Hamid Alvari Chenari Omid Pourheidari Ahmad Khodamipour	139
Board Characteristics, Audit Fees, and Political Connections: An Integration of Resource Dependence and Agency Theories	Ali Sayadi Iraj Mirhossini Reza Khajezadeh Saman Mirabbasi	163
The Relationship Between Managerial Overconfidence and ESG: Emphasizing The Role of Environmental Costing	Hanan Abdullah Hasan	187
Financial Stress Research Under the Magnitude Mind Lens	Taqi Al Abdwani Shima Ebrahimi Iliya Pishghadam	201

Editorial Board

Mohammad Reza Abbaszadeh

Associate Professor of Accounting, Faculty of Economics and Business Administration, Ferdowsi University of Mashhad, Mashhad, Iran

Masoud Azizkhani

Associate Professor of Accounting, Australian National University

Mohammad Ali Bagherpour Velashani

Associate Professor of Accounting, Faculty of Economics and Business Administration, Ferdowsi University of Mashhad, Mashhad, Iran

Emmanuel Constantine Mamatzakis

Professor of School of Business, Management and Economics London, United Kingdom

Alireza Dorestani

Associate Professor of Accounting, Northeastern Illinois University, College of Business and Management

Syed Zabid Hossain

Professor of Accounting and Information Systems, University of Rajshahi

Shokrollah Khajavi

Professor of Accounting, Tehran University, Tehran, Iran

Shahriar Khaksari

Professor of Finance, Suffolk University

Gholamhossein Mahdavi

Professor of Accounting, Shiraz University, Shiraz, Iran

Esfandiar Malekian

Professor of Accounting, University of Mazandaran, Mazandaran, Iran

Kaliyamoorthy Maran

Professor of Finance, Sri Sai Ram Engineering College, Chennai, Tamil Nadu, South India

Mahdi Moradi

Professor of Accounting, Faculty of Economics and Business Administration, Ferdowsi University of Mashhad, Mashhad, Iran

Mahmoud Mousavi Shiri

Associate Professor of Accounting, Payam Noor University, Tehran, Iran

Farzaneh Nasirzadeh

Associate Professor of Accounting, Faculty of Economics and Business Administration, Ferdowsi University of Mashhad, Mashhad, Iran

Balasundaram Nimalathasan

Professor of Accounting, Faculty of Management studies and commerce, University of Jaffna, Sir Lanka

Zabihollah Rezaee

Professor of Accounting, University of Memphis, USA

Ali Saeedi

Associate Professor of Accounting, University of Minnesota Crookston

Mahdi Salehi

Associate Professor of Accounting, Faculty of Economics and Business Administration, Ferdowsi University of Mashhad, Mashhad, Iran

Piotr Staszkievicz

Professor of Accounting, Warsaw, Poland

Kayhan Tajeddini

Professor of Strategic Management, Sheffield Hallam University, UK

Hassan Yazdifar

Professor of Accounting, Salford University Business School

Received articles will be published after professional referee, and being approved by the editorial board. Submission to this Journal is required to send at least two files include: an “Article cover” and an “Anonymous file of the main article”, through systems, <http://ijaaf.um.ac.ir/>. Submission of a manuscript implies that the paper has not been published before and it is not under consideration for publication elsewhere. Please pay attention to the following:

1-Article Cover

The file should be typed in Word software (office), and A4 sized page which include:

- Full title of the paper centered in Times New Roman 16.
- Author's name centered in Times New Roman 12 (Bold),
- Author's characteristics centered in Times New Roman 11 (italics), including the authors' name (the name of correspondence author should be marked with an asterisk), academic rank and the name of the institution or university (do not use pseudonym and nickname), complete address, telephone number and fax of correspondence author and e-mail address of all authors.

2- Anonymous file of the main article

The main article should be typed in Word software (office), A4 sized page and utmost in 7500 words (including references), and numbered pages while page margins are as follows: margin from the above 4 cm, from the below 5/6 cm, from the left 5/4 and from the right 5 cm. The line spacing should be set at 1 cm and the beginning of each paragraph indented about 0.3 cm, the entire text should be Justified and include the following:

The first page:

- Full title of the paper, centered in Times New Roman 16.
- Abstract in Times New Roman 11 and utmost 165 words, including four issues: (1) Subject and purpose of the paper, (2) methodology, (3) findings, and (4) conclusions, and Contribution (maximum).
- Utmost 5 keywords in Times New Roman
- Subject classification code in Times New Roman This coding is designed for subject classification in economic literature and how to use it in detail is available on the following website: www.aeaweb.org/journal/jel_class_system.html
- The main titles in manuscript (including abstract, introduction, literature review, methodology, findings and Conclusion) left aligned in Times New Roman 12(Bold) and sub-titles (such as Keywords, subject classification code and other necessary(important) titles) left aligned in Times New Roman 11(bold) and in the left alignment, content of main manuscript in Times New Roman
- As far as possible, do not use any image in the text and if used, it should be high- quality black and white photos.

Structure of second page until the end of manuscript is as follow:

- *Introduction* Some paragraphs contain explaining the problem, literature review, object (purpose), importance and necessity of it.
- *Literature review* a review of the literature investigates only related researches chronologically and the results exploit at the end of the section theory matrix or conceptual model that document research variables and Formulate research hypotheses.
- *Methodology* including Methods, data collection tools, population, sample size and sampling methods, analysis and model testing hypothesis, definition of study variables and operational definition of them can be in presented the same section that model testing is represented and there is no need to repeat.
- *Results* including the findings compare it with the findings of previous and interpretation of compliance or inconsistency of findings with research findings and theories. *Conclusion* includes a summary of the problem, provide a summary of the results and overall conclusion and

recommendations based on the results (policy recommendations is necessary only in applied research and, if necessary, recommendations for future research accordant with the research limitations or how development of current research;

- *References* are as Section 3-2 and
- *Endnotes* terminology and some necessary explanations provide in the endnote (no footer) and as a two-column table (contains the endnote number and content of endnote with invisible lines (No Border). Numbering the endnotes are as text and doing without the use of EndNote technique.

3- Other important points in the original file

3-1- *References inside the text*

- In order to reference within the text (APA) method should be used; so the author's last name with the year of publication it is presented in the text respectively. If there are two authors, last names' will be separated by "and" and if more than three people, "et al." will be used
- After the last name of the first author. If the number of resources is more than one, a semicolon (;) will be used to separate them.
- Any reference which is mentioned in the article, must be included in references part.
- If it is required to explain more about terms or words within the text, endnote can be used.

3-2- *References*

To set the reference list, use the (APA) method, as follows:

Books: Author, A.A. (Year of Publication). *Title of work*. Publisher City, State: Publisher.

When citing a book in APA, keep in mind:

- Capitalize the first letter of the first word of the title and any subtitles, as well as the first letter of any proper nouns.
- The full title of the book, including any subtitles, should be stated and *italicized*.
- Magazines: Author, A.A. (Year of Publication). Title of the Article. *Journal Title*, Volume (Issue), pp.-pp.
- When citing a magazine in APA, keep in mind:
- You can find the volume number with the other publication information of the magazine.
- You can typically find page numbers at the bottom corners of a magazine article.
- If you cannot locate an issue number, simply don't include it in the citation.
- Online magazine article: Author, A.A. (Year). Title of the Article. *Magazine Title*, Volume (Issue), Retrieved from <http://xxxx>
- When creating an online magazine citation, keep in mind:
- The volume and issue number aren't always on the same page as the article. Check out the other parts of the website before leaving it out of the citation.
- List of references should not be numbered. If there are 2 or more authors alike, in addition to alphabetical order, the publisher year also should be ordered. In order to avoid the pitfalls of the various sources listed, the beginning of each reference is 0/5 cm protrusion.
- If there are two authors, last names' will be separated by "and" and if more than three people, last names' will be separated by a comma (,) and the last one with "and".

3-3- *Charts, Tables and Formulas*

Charts' title should be written below and tables' title on the top centered in Times New Roman 11 and Bold. Charts and tables render within the text and immediately place after the explaining paragraph. The content of charts and tables in Times New Roman 10

- Tables and images number from No. (1) To refer to the content of the images and diagrams in the text, use the number and appropriate referral.
- Inserting decimal numbers, instead of using slash (/), and show negative numbers by using the minus sign.
- Tables include variables, should contain a column for the symbols used for variables.

- Formulas represent centered in Times New Roman¹¹ and render in a two column table with no border and number with figure in parentheses.

3-4- Other notes

- There is no need to spaces before dot, comma, semicolon, colon, question mark, exclamation mark and, but after these signs shall be placed a spaces.
- Use spaces before opening and after the closing parenthesis.
- Articles should not be sent to Persian/English journal or another language at home and/or abroad simultaneously.
- If the format and structure of submitted articles are not same as what mentioned in this guide, it won't be accepted.
- Journal can edit and delete some sections, without any change in its content and received journals will not bring.
- Responsibility for the accuracy of the article is the author.
- The file should be named in English. This name should include the first author's last name and sending date. Articles will be received only by Journal Website.
- In order to expedite the referee process and publishing an article, we will ask respected professors and researchers to ensure technical and literary editing and enforce this guide.

The journal is devoted to research papers. Research paper is extracted from the research project, master's dissertation or thesis.

Receiving the papers is only possible electronically.

Editor's Note

I am pleased to announce that the Ferdowsi University of Mashhad is publishing Iranian Journal of Accounting, Auditing & Finance (IJAAF). On behalf of the board of the IJAAF and my co-editors, I am glad to present the Volume 1, Issue 1 of the journal in December 2017; the journal will publish four issues in a year. The board includes experts in the fields of accounting, finance and auditing, all of whom have proven track records of achievement in their respective disciplines. Covering various fields of accounting, *IJAAF* publishes research papers, review papers and practitioner oriented articles that address significant issues as well as those that focus on Asia in particular. Coverage includes but is not limited to:

- Financial accounting
- Managerial accounting
- Auditing
- Taxation
- Accounting information systems
- Accounting education

Perspectives or viewpoints arising from regional, national or international focus, a private or public sector information need, or a market-perspective are greatly welcomed. Manuscripts that present viewpoints should address issues of wide interest among accounting scholars internationally and those in Asia in particular.

Yours faithfully,
Mahdi Moradi
Editor in Chief



RESEARCH ARTICLE

Sectoral Herding and Volatility in the Chinese Stock Market: Lessons from COVID-19

Syed Faisal Hassan Bukhari*, Pedro Pimentel, Gualter Couto

School of Business and Economics and CEEAplA, University of Azores, Ponta Delgada, Portugal

How to cite this article:

Bukhari, S. F. Hassan , Pimentel, P. and Couto, G. (2026). Sectoral Herding and Volatility in the Chinese Stock Market: Lessons from COVID-19. Iranian Journal of Accounting, Auditing and Finance, 10(3), 1-19. doi: 10.22067/ijaaf.2026.48161.1632
https://ijaaf.um.ac.ir/article_48161.html

ARTICLE INFO

Article History

Received: 2025-06-29

Accepted: 2025-11-22

Published online: 2026-07-12

Abstract

This study investigates the Chinese stock market's herd mentality and volatility both before and after the COVID-19 outbreak. Using the GARCH (1,1) and return dispersion (CSAD) models, the study examined how market volatility and investor behavior differ across economic sectors. Due to the market meltdown and instability caused by the COVID-19 pandemic, the results showed a significant rise in herd behavior and volatility following the outbreak. To control cross-sector herding and volatility in times of crisis, the research shows that sector-specific variations in these phenomena are necessary. The study highlights shortcomings, such as failing to consider more comprehensive factors like regional and global influences, and recommends that these areas be addressed in subsequent studies.

Keywords:

COVID-19, Diversification,
Efficient, Herding, Market
Hypothesis, Volatility

<https://doi.org/10.22067/ijaaf.2026.48161.1632>



NUMBER OF REFERENCES
84



NUMBER OF FIGURES
2



NUMBER OF TABLES
8

Homepage: <https://ijaaf.um.ac.ir>
E-ISSN: 2717-4131
P-ISSN: 2588-6142

*Corresponding Author:
Syed Faisal Hassan Bukhari
Email: nadealishanawer@gmail.com
Tel:
ORCID: 0000-0002-6971-1498

1. Introduction

The Chinese stock market was the first market to experience unprecedented economic uncertainty and havoc after the COVID-19 outbreak. The economy is continually struggling with the effects of the ongoing pandemic, trying to recover from fear and damage. Economic growth was significantly affected by containment and restrictions. After a slowdown to 2.6% in 2022, China experienced a rebound in economic growth, with an increase of 5.2% in 2023 and 5.0% in 2024. In the context of a financial crisis, the Chinese stock market experienced a notable decline, and return volatility increased during the pandemic's initial phase.

Due to fear and instability, investors' confidence has been adversely affected. This highlights a behavioral finance perspective: psychological biases, such as copying other decisions rather than fundamentals, are a primary source of herd behavior and market volatility. As irrational investor behavior intensifies during financial crises, the examination of herd behavior becomes more important. Consequently, investors need new strategies for portfolio diversification and hedging purposes ([Assaf et al., 2023](#)). The financial crisis resulted in price inefficiency, extreme volatility, and unpredictability; furthermore, the spread of herd behavior among investors during the crisis was lightning-fast ([Demirer et al., 2019](#); [IMF, 2023](#)).

While numerous studies exist on herding and volatility, there is currently no consensus regarding their origins, intensity, or similarities, and, to the best of our knowledge, there is a lack of sectoral analyses of volatile markets before and after COVID-19. There are different characteristics of COVID-19 compared to previous crises, essentially stemming from a compulsory hibernation of productive capacity in many economic sectors that remained mostly intact, waiting to recover after lockdowns. This paper explores herding and volatility at the sectoral level to identify driving factors and common features of herding and volatility before and after COVID-19. A method is presented to assess the impact of COVID-19 on investor sentiment and the resulting excessive volatility in the country at the epicenter.

Investors herd when they follow the market trend rather than their own expectations ([Chauhan et al., 2020](#)). Investors' uniform behavior is the result of similar beliefs and ideas widespread in the financial markets ([Ghorbel et al., 2022](#)). Investors' herd behavior causes returns to deviate from their fundamentals and, sooner or later, creates chaos in stock markets ([Javaira and Hassan, 2015](#)). The literature establishes that herd behavior intensity increases during adverse market conditions, thereby spiking volatility in the stock market. The COVID-19 pandemic boosted investors' irrational behavior and caused extreme volatility in the financial markets. All Chinese economic sectors have been affected, but with varying intensities. For instance, the real estate sector experienced a massive 20% decline, whereas the healthcare sector posted positive returns ([IMF, 2023](#)). Therefore, sectoral analysis is key to understanding the dynamics of the Chinese stock market after the COVID-19 outbreak.

Classical models, supported by research studies and risk-and-return theory, are unable to explain excessive price movements and herd behavior ([Wang et al., 2023](#)). Behavioral finance provides a theoretical foundation for understanding how irrational decisions and imitation can fuel market volatility, especially during crises. Behavioral finance has emerged as a field, claiming that excessive volatility results from many investors following others' decisions. During periods of chaos, it is not

uncommon for less informed investors to mimic the actions of other investors (Chang et al., 2020; Litimi et al., 2016). The financial markets have been significantly affected by the COVID-19 pandemic, leading to greater volatility and uncertainty. Investors are now relying more on general market sentiment rather than conducting their own individual analyses (Ampofo et al., 2023). Blasco et al. (2012) provide empirical evidence on the link between herding and volatility. Recent research discusses the COVID-19 financial crisis, with Ferreruela and Mallor (2021) claiming that existing herding and volatility vanish during the crisis and re-emerge afterwards.

This validates the behavioral finance theory that herd behavior is driven by market sentiment, which leads to temporary shifts in investor behavior (Espinosa-Méndez and Arias, 2021). COVID-19 has intensified herd behavior in financial markets as a short-term phenomenon. Therefore, this paper investigates whether herding is a short-term phenomenon and explores its origins, as well as those of volatility, across different economic sectors in the Chinese stock market.

In addition to a description of what happened during COVID-19 in terms of the financial crisis, what was special about COVID-19 that affected herding and volatility, and how different economic sectors were affected and how they faced its effects, relative to other crises, are worth knowing.

In this paper, no differences were found between the before- and after-COVID-19 periods across all economic sectors. Although all sectors exhibited a similar reaction, the prevalence of herding and volatility is more noticeable on a sector-by-sector basis than when considered in aggregate.

This observation indicates that crises stemming from productive capacity hibernation do not uniformly impact or impose a challenging slow recovery across economic sectors, as has occurred in other financial crises. Based on our current knowledge, this particular phenomenon was previously unknown. Our findings also revealed that herding and volatility are strongly present, which calls for effective reforms and improved information efficiency.

This study provides useful insights into herd behavior and volatility, offering new knowledge to the literature that could be generalizable through further empirical research in other countries and using data from similar COVID-19 financial crises. Research findings provide insight into a volatility component not associated with investors' behavior across all economic sectors, which is fruitful for forecasting procedures, hedging strategies, and portfolio diversification. The authors observe a significant lack of studies that focus on both volatility and herding behavior of investors, specifically in the context of sectoral versus aggregate analyses in the Chinese stock market. Consequently, this paper seeks to address this gap in the existing literature and advance our understanding of investor behavior across these distinct economic sectors.

The paper proceeds with part two for the literature review, part three for the data and methodology, part four for the analysis and results, and part five for the concluding remarks.

2. Theoretical foundations

2.1. Herding in equity markets

To address herd behavior, many theories have examined herd formation in financial markets. For instance, some describe herd behavior as most evident and damaging during the financial crisis (Chiang & Zheng, 2010). In contrast, others conclude that herd behavior intensifies during periods of lower uncertainty (Hudson et al., 2020). Herd behavior is difficult to define, and researchers have not

agreed on a single definition. [Welch \(1992\)](#) defines herding as “investors who imitate the actions of other investors intentionally and ignore their own private information”. Herding occurs when a group of investors reaches consensus, influenced by economic and social factors. Investors choose the herd when sufficient information is unavailable to construct an optimal portfolio, and their main objective is to protect their investments from uncertainties, market risk, and non-parametric results ([Lee et al., 2021](#)).

[Chang et al. \(2000\)](#) claimed that during the financial crisis, investors ignore the available private information and follow the market consensus. Due to herd behavior, individual investors' returns tend to move in tandem with market returns, leading to high volatility and making it difficult to reap the benefits of diversification. Investors' aggregate behavior influences other investors' decisions and contributes to market inefficiency. The literature on herding behavior explains several associated risks, like price volatility, asymmetric effects in returns, speculation, market bubbles, errors in stock prices, imbalance between market variables, and market crashes ([Sihombing et al., 2021](#)). Herding promotes volatility and disturbs market trading, requiring regulators and policymakers' immediate attention to stabilize the stock market ([Demirer and Kutan, 2006](#)).

Several studies have examined herd behavior in the Chinese stock market, but only after the COVID-19 pandemic. For instance, [Espinosa Méndez and Maquieira \(2023\)](#) claimed that the Chinese stock market experienced strong herd behavior after COVID-19, with herd behavior dominance during low volatility and high returns. [Chang et al. \(2000\)](#) noted herding when investors sold their oil stocks due to fear and uncertainty after the COVID-19 crisis. By contrast, [Wu et al. \(2020\)](#) concluded that herding intensity decreased after COVID-19. [Yuan \(2021\)](#) looked at herding at the sectoral and aggregate levels, but findings fail to generalize the presence of herding across economic sectors before COVID-19. So,

H1. COVID-19 influences herd behavior across economic sectors.

2.2. Volatility in equity markets

Asian stock markets experienced sharp declines and extreme volatility during the COVID-19 outbreak. A highly volatile market has weakened investors' confidence and created doubts ([Mishra and Mishra, 2021](#)). Fear of a devastating virus has increased the likelihood of abrupt price movements, market crashes, and financial crises. Significant empirical evidence of investors' fear and anxiety led to unprecedented volatility ([Vuong et al., 2022](#)). Stock market volatility is a barometer of risk; therefore, volatility is a key factor for stock market performance ([Zaremba et al., 2020](#)). Exploring the causes, timing, and magnitude of volatility is essential for understanding and navigating instability and financial turmoil. In addition, social media plays a pivotal role in spreading pandemic-related rumors and creating fear, thereby risking the financial system and destabilizing financial markets.

Some notable empirical studies have examined the effects of COVID-19 and how the pandemic led to greater uncertainty and significant setbacks for the Chinese economy. [Li et al. \(2022\)](#) stated that high trading activity significantly contributes to high volatility, causes substantial negative sentiment, and leads investors to take “flights to safety”. Research argues that excessive stock market

volatility results from the association between information asymmetry and investors' sentiment (Gao et al., 2022b; Li et al., 2022). Choi (2021) found that aggregate market index analysis yields homogeneous results and obscures critical information needed for future decisions. In addition, earlier studies have examined how volatility has passed through the stock market during health crises; however, no study has examined extreme volatility across economic sectors.

This paper also examines volatility in the Chinese stock market to anticipate potential threats to the financial system and its stability. The Chinese stock market is a promising and large emerging market, but it has imposed some restrictions on foreign investment and is described as a segmented market (Wang et al., 2022). Historically, the Chinese stock market has been vulnerable, generating highly volatile abnormal returns during economic crises (Li et al., 2022; Ramelli and Wagner, 2020; Wu et al., 2020). Identifying the drivers of stock market volatility after COVID-19 provides stakeholders with insights to address issues and control the further financial impact of pandemic crises. Previous research provides concrete evidence that fear and volatility across economic sectors are uneven. According to Mensi et al. (2017), the economic sectors of finance, technology, energy, and telecommunications are the largest contributors to stock market volatility. Zhang and Giouvris (2022) analyzed cross-sectional economic sectors' volatility before and after COVID-19, suggesting that volatility varies across economic sectors in the long run. Relying on the literature discussed above, the following research question arises.

H2. COVID-19 influences volatility across economic sectors.

3. Data and methodology

Daily closing prices from January 1, 2015, to December 31, 2024, for all companies listed on the Shanghai Stock Market were obtained from Thomson Reuters Data Stream and subdivided into 10 economic sectors according to the Thomson Reuters Business Classification (TRBC). TRBC is a useful tool for investors when selecting investment portfolios across developed and emerging markets. The pre-COVID-19 period was defined as January 2015 to December 2019, and the post-COVID-19 period as January 2020 to December 2024. The time series data, comprising both aggregate and sectoral components, were tested for stationarity using the Dickey–Fuller (ADF) unit root test.

3.1. Herd behavior detection

When the stock market is under stress, investors copy other investors' actions and overlook information available to them. Such behavior reduces the cross-sectional variation among the stocks and causes herding in the equity market (Christie and Huang, 1995). Cross-Sectional Standard Deviation (CSSD) model can check the presence of herd behavior in the stock market using the following equation for each economic sector/market:

$$CSSD_t = \sqrt{\frac{\sum_{i=1}^N (r_{i,t} - r_t)^2}{N-1}} \quad (1)$$

where $CSSD_t$ is the cross-sectional standard dispersion of returns among all companies in the same sector/market on a given day t , $r_{i,t}$ is the stock return for the company i on day t , r_t is the sector/market average stock return on a given day t , and N is the total number of companies in the

sector/market.

However, [Chang et al. \(2000\)](#) put forward a new model based on the Capital Asset Pricing Model (CAPM), denoted by followers as the CCK model. This model captures investors' herd behavior through a nonlinear association between Cross-Sectional Absolute Deviation (CSAD) and market returns. For example, when a large number of investors form a herd, returns cluster around a single point. However, if investors take an anti-herding stance, returns are dispersed ([Sheikh et al., 2025](#)). [Tan et al. \(2008\)](#) argue that the CCK model is better at detecting herd behavior than the CSSD model. Therefore, this paper adopts the CSAD model to check herd behavior in the Chinese stock market, using the following equation:

$$CSAD_t = \alpha + \gamma_1 |r_t| + \gamma_2 r_t^2 + e_t \quad (2)$$

with $CSAD_t$ given by the following equation:

$$CSAD_t = \frac{1}{N} \sum_{i=1}^N |r_{i,t} - r_t| \quad (3)$$

where $CSAD_t$ is the mean of the cross-sectional absolute return of all companies in a sector/market, r_t is the sector/market average stock return on a given day t , $r_{i,t}$ is the stock return for the company i on day t , N is the total number of companies in the sector/market, α is a constant, the bottom cross-sectional absolute return of all companies in a sector/market, γ_1 and γ_2 are coefficients to be estimated.

The coefficients γ_1 and γ_2 represent the linear and nonlinear relationships, respectively, between the sectoral return and the cross-sectional absolute return of all companies within a sector. If γ_2 is negative and statistically significant, indicating that herd behavior exists in that economic sector/market. During periods of financial crisis and uncertainty, large price movements should occur in the stock market, thereby boosting herd behavior. Similarly, unequal rates of fluctuation between $CSAD$ and sector/market returns show a severe presence of herd behavior ([Chang et al., 2000](#)).

3.2. GARCH(1,1) Effect

The autoregressive conditional heteroskedasticity (GARCH (1,1)) model is employed to capture volatility in the Chinese stock market before and after the COVID-19 pandemic. The GARCH(1,1) model is used to forecast and assess volatility in time series. EGARCH and TGARCH models are used to capture persistent volatility and the leverage effect. Thus, the GARCH(1,1) model performs better at capturing sudden volatility clustering, which is vital for the before-and-after COVID-19 evaluation. The so-called random coefficient problem in the GARCH (1,1) model can be addressed by calculating variations across different sample periods. The baseline idea of the GARCH (1,1) model is “large price movements followed by large price movements and low price movements followed by low price movements”. To check the presence of volatility in the stock market in the aggregate and sectoral context, the following GARCH (1,1) model was set:

Conditional Mean Eq

$$R_t = \alpha_0 + \alpha_1 R_{t-1} + \varepsilon_t \quad (4)$$

Where: $\varepsilon_t \sim N(0, \sigma_t^2)$

Conditional Mean Eq

$$\sigma_t^2 = \beta_0 + \beta_1 \mu_{t-1}^2 \quad (5)$$

In the conditional mean equation, R_t are the returns of stocks, as per the Efficient Market

Hypothesis (EMH), R_t has mean-reversion behavior; therefore, returns are highly volatile and unpredictable. Equation No.5 represents the conditional variance equation, where σ_t^2 depends on the squared lag values σ_{t-1}^2 and μ_{t-1}^2 is the GARCH (1,1) term. Investors' financial decision-making depends on risk and returns; the GARCH(1,1) model is used to estimate risk (volatility) and return, with risk (volatility) derived from the mean equation and return from the variance equation. Therefore, this study employed the GARCH (1,1) model to assess volatility across all stock market sectors during both stable and financial crisis periods. The AR(1) model is selected to analyze the conditional mean equation based on AIC and BIC; it is simple and fits well across all sectors.

4. Empirical analysis and discussion

4.1. Descriptive statistics

Descriptive statistics of r_t and $CSAD_t$ for each sector and at the aggregate market level, before and after COVID-19, are presented in Tables 1 and 2. The descriptive statistics of the CSAD and sectoral returns before and after COVID demonstrate persistent distributional patterns. In both sub-periods, most sectoral returns exhibit negative skewness, suggesting an asymmetric distribution and a greater likelihood of negative returns. The values of the Kurtosis coefficients exceed 3 in most sectors (except healthcare), indicating leptokurtic distributions; these results are well documented in the financial literature.

Comparison of the sub-periods reveals that the post-COVID period exhibits higher mean returns and greater cross-sectional dispersion, particularly in the telecom and technology sectors. This suggests that the stock market faces greater uncertainty and risk following the pandemic shock. The ongoing presence of Non-normal return distributions in both sub-periods supports the use of the CSAD to capture herding behavior, as traditional models assume a normal distribution and cannot detect underlying return behavior. The minimum CSAD values are reported as zero due to rounding to three decimal places.

Table 1. Descriptive statistics before COVID-19

Sector	Variable	Mean	Median	SD	Kurt.	Skew.	Min	Max
Cyclical	r_t^{cc}	-0.025	0.002	0.563	5.573	-1.218	-3.389	2.026
Consumer	$CSAD_t^{cc}$	0.548	0.545	0.202	2.904	-0.453	0.000	1.351
Energy	r_t^{en}	-0.026	0.000	0.638	4.286	-1.124	-3.422	1.990
	$CSAD_t^{en}$	0.575	0.545	0.289	1.035	0.514	0.000	1.709
Financials	r_t^{fin}	-0.005	0.000	0.409	6.896	-0.514	-2.469	2.368
	$CSAD_t^{fi}$	0.405	0.382	0.188	3.867	0.856	0.000	1.415
Healthcare	r_t^{hc}	-0.011	0.000	0.571	3.150	-0.655	-3.100	2.167
	$CSAD_t^{hc}$	0.559	0.552	0.225	1.871	-0.175	0.000	1.537
Industrials	r_t^{in}	-0.019	0.003	0.576	5.590	-1.203	-3.458	2.169
	$CSAD_t^{in}$	0.551	0.545	0.215	2.303	-0.202	0.000	1.411
Materials	r_t^{ma}	-0.018	0.000	0.572	4.672	-1.066	-3.265	2.074
	$CSAD_t^{ma}$	0.551	0.544	0.212	2.554	-0.235	0.000	1.357
Non-Cyclical	r_t^{nc}	-0.019	0.000	0.548	4.089	-0.853	-3.263	2.116
	$CSAD_t^{nc}$	0.566	0.554	0.234	3.080	0.209	0.000	1.811
Technology	r_t^{it}	-0.003	0.000	0.741	3.243	-0.722	-3.931	2.397
	$CSAD_t^{it}$	0.661	0.653	0.265	2.335	-0.081	0.000	1.912
Telecom	r_t^{tc}	0.041	0.029	0.956	1.602	-0.486	-4.360	2.749
	$CSAD_t^{tc}$	0.697	0.541	0.551	0.402	0.949	0.000	2.773
Utilities	r_t^{ut}	-0.016	0.000	0.499	6.346	-1.224	-2.986	1.899
	$CSAD_t^{ut}$	0.450	0.439	0.205	1.836	0.388	0.000	1.267
Aggregate	r_t	-0.016	0.003	0.532	5.452	-1.085	-3.158	2.079
	$CSAD_t$	0.548	0.541	0.204	3.576	-0.220	0.000	1.429

Table 2. Descriptive statistics after COVID-19

Sector	Variable	Mean	Median	SD	Kurt.	Skew.	Min	Max
Cyclical	r_t^{cc}	0.005	0.000	0.581	7.391	-1.343	-4.600	2.117
Consumer	$CSAD_t^{cc}$	0.714	0.740	0.238	3.758	-1.483	0.000	1.787
Energy	r_t^{en}	0.014	0.000	0.676	4.077	-0.760	-4.672	2.043
	$CSAD_t^{en}$	0.709	0.698	0.324	0.813	0.027	0.000	1.950
Financials	r_t^{fin}	-0.013	0.000	0.543	6.796	-0.662	-4.240	2.926
	$CSAD_t^{fi}$	0.522	0.517	0.216	1.559	-0.156	0.000	1.235
Healthcare	r_t^{hc}	0.003	0.000	0.650	1.740	-0.288	-2.958	2.581
	$CSAD_t^{hc}$	0.733	0.743	0.299	10.577	0.477	0.000	3.538
Industrials	r_t^{in}	0.007	0.023	0.572	6.919	-1.283	-4.418	1.880
	$CSAD_t^{in}$	0.656	0.686	0.233	2.317	-1.208	0.000	1.443
Materials	r_t^{ma}	0.016	0.012	0.598	5.303	-1.081	-4.288	1.995
	$CSAD_t^{ma}$	0.721	0.751	0.268	1.699	-0.928	0.000	1.495
Non-Cyclical	r_t^{nc}	0.007	0.000	0.572	5.120	-1.018	-3.851	1.858
	$CSAD_t^{nc}$	0.708	0.721	0.274	1.911	-0.563	0.000	1.837
Technology	r_t^{it}	-0.009	0.000	0.778	3.898	-0.695	-5.122	2.303
	$CSAD_t^{it}$	0.757	0.758	0.288	5.427	-0.234	0.000	2.887
Telecom	r_t^{tc}	0.247	0.255	0.783	3.330	-0.719	-4.571	2.681
	$CSAD_t^{tc}$	0.891	0.933	0.472	-0.505	0.167	0.000	2.391
Utilities	r_t^{ut}	0.006	0.000	0.618	4.478	-0.762	-4.227	2.334
	$CSAD_t^{ut}$	0.547	0.535	0.260	1.512	0.357	0.000	1.732
Aggregate	r_t	0.004	0.003	0.550	6.589	-1.249	-4.158	1.748
	$CSAD_t$	0.713	0.747	0.238	3.694	-1.506	0.000	1.771

1303 observations

During the COVID-19 period, the financial sector had the lowest mean return, while the telecom sector had the highest. A similar pattern is observed for CSAD mean return: the financial sector has the lowest, and the telecom sector the highest. Risk is estimated using the standard deviation; the telecom sector has the highest risk among all sectors, indicating it is more volatile than other sectors of the Chinese stock market ([Chiang and Zheng, 2010](#)). The skewness values indicate that the Gaussian distribution is not present in all sectors, and the returns of all sectors are left-skewed. Kurtosis coefficient values for all sectoral returns are higher than 3, except for healthcare, indicating that sectoral returns are leptokurtic, with a thick-tailed distribution.

During the COVID-19 period, the financial sector had the lowest mean return, while the telecom sector had the highest. A similar pattern is observed for CSAD mean return: the financial sector has the lowest mean return, and the telecom sector has the highest. Risk is estimated using standard deviation; the telecom sector has the highest level of risk among all sectors, indicating that it is more volatile than other sectors of the Chinese stock market (Chiang and Zheng, 2010). The skewness values indicate that the Gaussian distribution is not present in all sectors, and the returns of all sectors are left-skewed. Kurtosis coefficient values for all sectoral returns are higher than 3 except for healthcare; this indicates that sectoral returns have a leptokurtic character, meaning the tails of the distribution are thicker.

4.2. Herd behavior

Tables 3 and 4 present the sectoral and aggregate regression results of the CSAD model to capture herd behavior in the Chinese stock market. The R-squared values for all sectoral and aggregate regressions are acceptable and vary across sectors. The γ_2 coefficient is significant and negative for all sectors except the healthcare sector and aggregate data sets before the COVID period. The healthcare and aggregate data sets are positive and significant, which indicates reverse herding. During reverse herding, investors make irrational decisions, ignore fundamental values, and increase return dispersion. However, the γ_2 coefficient demonstrates that the nonlinear association between CSAD and squared mean return is present in all sector samples after the COVID-19 pandemic. The negative sign indicates that herd behavior is present in the stock market, and the CSAD coefficient decreases rather than increases, corresponding to a high value of r_t^2 . The results confirm our research question Q1: herd behavior is spotted in all sectors and aggregate data sets, except for healthcare and aggregate data sets before the COVID-19 period. The findings of this study are consistent with those of Chang et al. (2000), Espinosa Méndez and Maquieira (2022), Sheikh et al. (2025), Wu et al. (2020), and Yuan (2021).

Table 3. Herd behavior – sectoral and aggregate before COVID-19

Sectors	γ_1	γ_2	α	R^2
Materials	0.388*** (0.031)	-0.055*** (0.015)	0.417*** (0.010)	0.328
Cyclical Consumer	0.370*** (0.029)	-0.052*** (0.013)	0.425*** (0.009)	0.329
Financials	0.544*** (0.030)	-0.070*** (0.019)	0.271*** (0.007)	0.537
HealthCare	-0.050*** (0.013)	0.118*** (0.010)	0.520*** (0.007)	0.199
Industrials	0.356*** (0.031)	-0.046*** (0.014)	0.429*** (0.010)	0.297
Non-Cyclical	0.430*** (0.036)	-0.059*** (0.019)	0.420*** (0.011)	0.321
Technology	0.427*** (0.033)	-0.083*** (0.013)	0.483*** (0.013)	0.278
Utilities	0.408*** (0.032)	-0.049*** (0.016)	0.328*** (0.009)	0.360
Energy	0.473*** (0.040)	-0.068*** (0.017)	0.393*** (0.014)	0.301
Telecom	0.882*** (0.061)	-0.203*** (0.023)	0.264*** (0.030)	0.289
Aggregate	-0.102*** (0.012)	0.099*** (0.008)	0.519*** (0.006)	0.298

***, ** and * denote significance at <0.01, <0.05 and <0.1, respectively. Standard deviation in parentheses.

This model explains that herding occurs in the market when investors follow the average market opinion. These findings confirm the presence of the Bandwagon Effect (tendency to rally on public opinion) in all sectors of the market and violate the assumptions of the efficient market hypothesis by indicating anomalies in the stock market during the period of crisis (Lee et al., 2021). These results also confirm that during the period of COVID-19, investors feel confident by copying the actions of other investors and reduce uncertainty (Mishra and Mishra, 2023). The empirical findings imply that COVID-19 has generated a market bubble that stimulates cross-sector herding in the stock market. Resultantly, herd behavior of investors caused extreme volatility in all sectors of the market and affected the returns of optimal portfolios (Bukhari et al., 2021). During the period of chaos and extreme uncertainty, investors panicked and were forced towards herd behavior instead of following their own beliefs (Sihombing et al., 2021). In this situation, market efficiency is compromised, and investors will not be able to achieve desired goals by portfolio diversification across sectors (Bukhari et al., 2022). The panic and irrational behavior of investors create a vicious cycle that spreads across all sectors of the market.

Table 4. Herd behavior – sectoral and aggregate after COVID-19

Sectors	γ_1	γ_2	α	R^2
Materials	0.485*** (0.037)	-0.129*** (0.016)	0.562*** (0.014)	0.205
Cyclical Consumer	0.412*** (0.037)	-0.109*** (0.013)	0.584*** (0.012)	0.181
Financials	0.460*** (0.028)	-0.099*** (0.013)	0.383*** (0.009)	0.318
HealthCare	0.600*** (0.047)	-0.163*** (0.025)	0.512*** (0.016)	0.266
Industrials	0.429*** (0.032)	-0.108*** (0.013)	0.522*** (0.012)	0.210
Non-Cyclical	0.451*** (0.041)	-0.104*** (0.019)	0.559*** (0.014)	0.180
Technology	0.355*** (0.031)	-0.045*** (0.011)	0.590*** (0.014)	0.244
Utilities	0.499*** (0.034)	-0.112*** (0.015)	0.376*** (0.012)	0.293
Energy	0.505*** (0.041)	-0.122*** (0.017)	0.518*** (0.017)	0.191
Telecom	0.639*** (0.059)	-0.175*** (0.023)	0.612*** (0.028)	0.145
Aggregate	0.419*** (0.035)	-0.095*** (0.015)	0.581*** (0.012)	0.196

***, ** and * denote significance at <0.01, <0.05 and <0.1, respectively. Standard errors in parentheses

The study has highlighted uncorrelated sectors in the stock market, which are useful for investors when selecting a balanced portfolio (basket). Further, the findings of this study provide useful insights for all stakeholders to take necessary remedial measures to overcome the negative effects of herd behavior during financial turmoil.

4.3. Volatility

Results for volatility across sectors and the aggregate, before and after COVID-19, are presented in Tables 5 and 6 and Figures 1 and 2. Before estimating the GARCH (1,1) effect, we check

stationarity in all data streams and employ the Augmented Dickey-Fuller (ADF) test. Results show that all data streams are stationary at the level, and the null hypothesis is rejected. The GARCH (1,1) effect indicates that volatility is present in only 3 sectors: healthcare, telecom, and utilities; all other sectors and aggregate data streams show stability before COVID-19. Our results confirm that the magnitude of volatility varies across the sectors and the aggregate market sample. However, after COVID-19, returns remain highly volatile, and the GARCH(1,1) effect is evident across all sectoral and aggregate data sets. COVID-19 is highly harmful to all sectors, and volatility is more pronounced afterwards. This supports our research question that the magnitude of volatility differs across sectors and aggregate data streams. Study findings are consistent with previous studies that volatility significantly increased during the period of COVID-19 (Gao et al., 2022a; Zhang and Giouvris, 2022).

Table 5. Volatility – sectoral and aggregate data sets (Before COVID-19)

Sectors	β_0	β_1	ADF
Materials	-0.00558 (0.033)	0.319*** (0.011)	-28.134
Cyclical Consumer	0.0014 (0.026)	0.406*** (0.015)	-26.485
Financials	-0.0109 (0.024)	0.169*** (0.005)	-27.436
HealthCare	0.153*** (0.050)	0.282*** (0.012)	-27.974
Industrials	-0.0137 (0.031)	0.336*** (0.011)	-27.353
Non-Cyclical	-0.0137 (0.020)	0.331*** (0.011)	-28.220
Technology	0.0406 (0.035)	0.288*** (0.010)	-28.490
Utilities	0.0364 (0.035)	0.529*** (0.021)	-27.888
Energy	0.283*** (0.062)	0.685*** (0.042)	-26.330
Telecom	0.204*** (0.041)	0.206*** (0.007)	-27.17
Aggregate	-0.0101 (0.027)	0.286*** (0.009)	-28.062

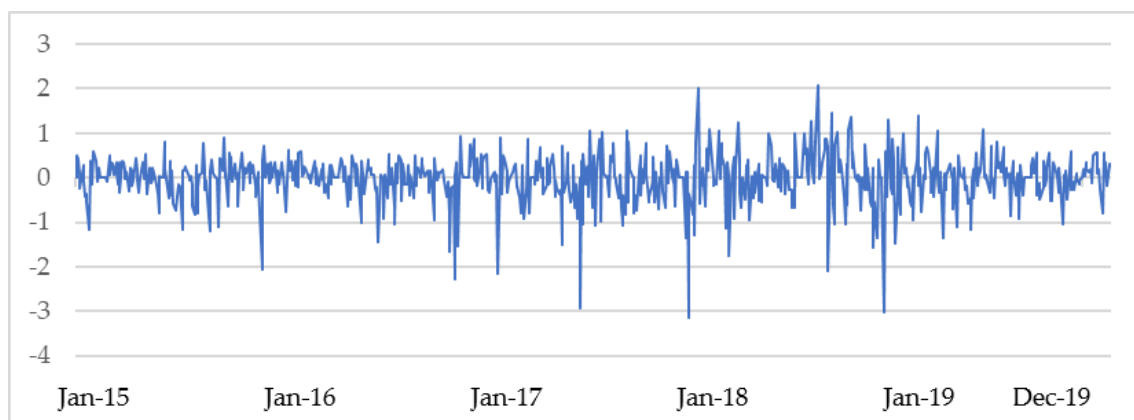
***, ** and * denote significance at <0.01, <0.05 and <0.1, respectively.
Standard errors in parentheses.

Table 6. Volatility – sectoral and aggregate data sets (After COVID-19)

Sectors	β_0	β_1	ADF
Materials	0.182*** (0.044)	0.280*** (0.015)	-27.070
Cyclical Consumer	0.238*** (0.052)	0.356*** (0.022)	-28.658
Financials	0.257*** (0.050)	0.232*** (0.013)	-26.962
HealthCare	0.298*** (0.058)	0.312*** (0.023)	-22.237
Industrials	0.294*** (0.048)	0.243*** (0.014)	-27.044
Non-Cyclical	0.184*** (0.046)	0.297*** (0.016)	-26.089
Technology	0.118*** (0.041)	0.291*** (0.014)	-25.296
Utilities	0.122*** (0.043)	0.530*** (0.027)	-28.406
Energy	0.118** (0.046)	0.541*** (0.028)	-22.356
Telecom	0.303*** (0.049)	0.277*** (0.015)	-27.419
Aggregate	0.208*** (0.044)	0.245*** (0.013)	-27.025

***, ** and * denote significance at <0.01, <0.05 and <0.1, respectively.

The results of the COVID-19 period are understandable, as all sectors were affected by an unprecedented pandemic that disrupted market trading, and sectoral volatility is strongly associated with the COVID-19 outbreak. Chinese financial market faced serious financial turmoil and remained sensitive after the outbreak of the pandemic. About 5 years of data (from Jan 2020 to Dec 2024) for the Chinese stock market show the far-reaching impact of COVID-19 on sectoral and aggregate returns. Our findings also suggest that investor sentiment has produced more sectoral volatility, especially after COVID-19 (Gao et al., 2022a). The results provide a clear picture to investors that risk has increased due to extreme volatility and herd behavior in the Chinese stock market after the pandemic (Bukhari et al., 2021; Gao et al., 2022b). Based on the paper's findings, investors can develop new investment strategies to hedge against risks posed by the COVID-19 pandemic across all sectors of the stock market.

**Figure 1.** Aggregate returns before COVID-19

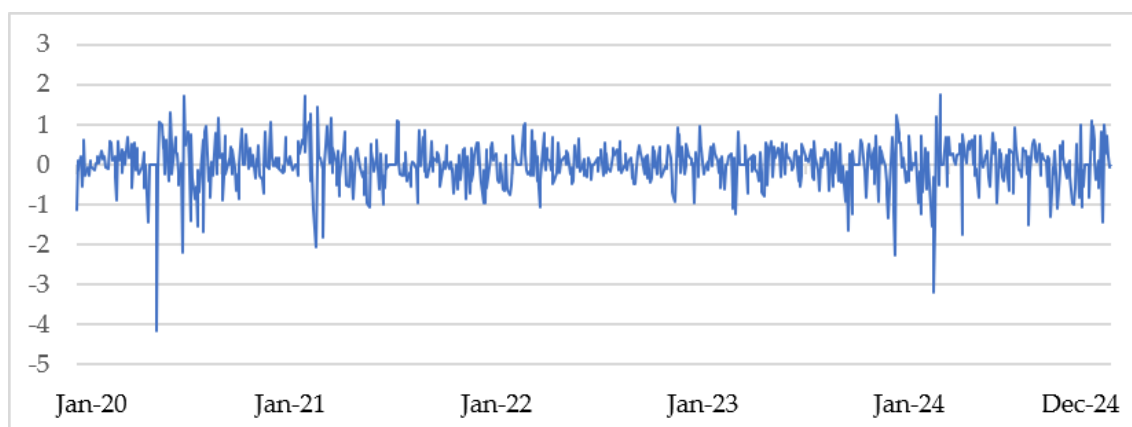


Figure 2. Aggregate returns after COVID-19

The outcomes of this study provide a comprehensive understanding of the root causes of panic and market chaos after COVID-19. The effects of volatility can be minimized by increasing information efficiency for investors during periods of chaos. These findings are valuable for the participants regarding expected future crises in the market/sector. After analyzing these results, fund managers can explore the causes and effects of sector-level volatility to develop comprehensive investment strategies that deliver productive results.

4.4. Robustness test

To ensure the robustness of our findings, we split the data into 10-year windows (2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, and 2024) to verify the presence of herd behavior and volatility. The results for each year are calculated to assess robustness and are presented in Tables 7 and 8. The overall findings from robust testing are not substantially different from the results already reported in (3-6). Chinese stock market experienced extreme volatility and herd behavior after the COVID-19 outbreak. This corroborates our main empirical findings and supports the reliability of our results.

Table 7. Robustness Testing – Herd Behavior

Years	γ_1	γ_2	α	R^2
2015	-0.119*** (0.024)	0.159*** (0.018)	0.498*** (0.010)	0.409
2016	-0.125*** (0.027)	0.210*** (0.023)	0.510*** (0.011)	0.417
2017	-0.133*** (0.028)	0.242*** (0.030)	0.489*** (0.010)	0.421
2018	-0.101*** (0.020)	0.082*** (0.012)	0.527*** (0.011)	0.358
2019	-0.070*** (0.020)	0.097*** (0.014)	0.525*** (0.012)	0.213
2020	0.159*** (0.035)	-0.028*** (0.011)	0.309*** (0.014)	0.198
2021	0.330*** (0.049)	-0.056*** (0.019)	0.555*** (0.018)	0.186
2022	0.509*** (0.066)	-0.135*** (0.031)	0.539*** (0.024)	0.286
2023	0.561*** (0.072)	-0.188*** (0.037)	0.595*** (0.028)	0.243
2024	0.574*** (0.075)	0.193*** (0.039)	0.598*** (0.029)	0.301

***, ** and * denote significance at <0.01, <0.05 and <0.1, respectively. Standard errors in parentheses.

Table 8. Robustness testing – Volatility

Years	β_0	β_1	ADF
2015	0.141* (0.097)	0.346*** (0.022)	-27.646
2016	-0.0119* (0.077)	0.440*** (0.028)	-26.835
2017	0.137* (0.081)	0.134*** (0.012)	-27.747
2018	-0.032 (0.044)	0.386*** (0.021)	-28.850
2019	1.071*** (0.125)	0.116*** (0.016)	-26.325
2020	-0.021 (0.046)	0.367*** (0.022)	-27.856
2021	-0.031 (0.084)	0.281*** (0.032)	-28.732
2022	0.214** (0.097)	0.310*** (0.039)	-27.585
2023	0.235** (0.098)	0.337*** (0.042)	-28.525
2024	0.215 (0.088)	0.324*** (0.040)	-27.805

***, ** and * denote significance at <0.01, <0.05 and <0.1, respectively. Standard errors in parentheses.

5. Conclusion

This paper investigated sectoral herding and volatility in the Chinese stock market, using data from Jan 2015 to Dec 2024 (before and after the pandemic). We employed [Chang et al. \(2000\)](#) model to check for herd behavior and an AR (1) was used to examine volatility. [Li et al. \(2022\)](#) claim that AR (1) is reliable in capturing the volatility and fear among the participants of the stock market. The impact of the pandemic was massive, and the stock market's sudden crash triggered fear and unprecedented instability. Our results showed that cross-sectoral herding and volatility have increased during the pandemic. The Chinese stock market was highly sensitive during COVID-19 compared with the pre-COVID-19 period and experienced excessive volatility. However, sectoral and aggregate herding and volatility exhibit distinct features across market conditions. It provides an opportunity for all investors to select a balanced portfolio across sectors when a COVID-19 crisis arises. Nevertheless, investors should not pursue high returns during excessive volatility and herding ([Wu et al., 2020](#)).

This paper contributes to the literature in several ways, as it examines herding and volatility at the sectoral and aggregate levels under different market conditions, providing useful insights into optimal portfolio selection. The results showed that herding activity is stronger during periods of high volatility; however, low volatility has not weakened herd behavior.

The robustness of the results also confirms that the stock market was semi-stable and efficient before COVID-19; however, the pandemic has had a devastating effect on the Chinese stock market. As a result, policymakers must be more concerned about uncertainty and instability in the market. Regulators should intervene and support advanced systems that can detect volatility clustering and enable pre-emptive measures against herd behavior. In addition, policymakers should develop information transparency requirements so that all listed companies disclose relevant information during periods of crisis and uncertainty. Timely and accurate information prevents irrational decision-making and reduces herding and market trends. Cross-sector herding and volatility during both periods, which require comprehensive strategies to mitigate the effects of future crises alike.

The main research limitation was that herding is a complex phenomenon, and it is difficult to disassociate herding from other investors' behavioral patterns. Future studies can examine the cross-sector contagion effects of volatility and herding across Asian stock markets, incorporating macroeconomic variables and diagnostic tests.

References

1. Ampofo, R. T., Aidoo, E. N., Ntiamoah, B. O., Frimpong, O. and Sasu, D. (2023). An empirical investigation of COVID-19 effects on herding behaviour in USA and UK stock markets using a quantile regression approach. *Journal of Economics and Finance*, 47(2), pp. 517-540. <https://doi.org/10.1007/s12197-022-09613-8>
2. Assaf, A., Mokni, K., Yousaf, I. and Bhandari, A. (2023). Long memory in the high frequency cryptocurrency markets using fractal connectivity analysis: The impact of COVID-19. *Research in International Business and Finance*, 64(1), A. 101821. <https://doi.org/10.1016/j.ribaf.2022.101821>
3. Blasco, N., Corredor, P. and Ferreruela, S. (2012). Does herding affect volatility? Implications for the Spanish stock market. *Quantitative Finance*, 12(2), pp. 311-327. <https://doi.org/10.1080/14697688.2010.516766>
4. Bukhari, F. H., Ahmad, H. and Hanif, H. (2021). Sectoral herding and contagion effect: pre and during COVID-19 crisis. *VFAST Transactions on Education and Social Sciences*, 9(4), pp. 143-152. <https://doi.org/10.21015/vtess.v9i4.792>
5. Bukhari, S. F. H., Ahmad, H., Hanif, H. and Shah, S. F. (2022). Volatility contagion (spillover) between Chinese and Pakistani stock markets during COVID-19: pre and post analysis of trade-level data. *International Journal of Monetary Economics and Finance*, 15(4), pp. 309-330. <https://doi.org/10.1504/IJMEF.2022.127157>
6. Chang, C. L., McAleer, M. and Wang, Y. A. (2020). Herding behaviour in energy stock markets during the Global Financial Crisis, SARS, and ongoing COVID-19. *Renewable and Sustainable Energy Reviews*, 134, A. 110349. <https://doi.org/10.1016/j.rser.2020.110349>
7. Chang, E. C., Cheng, J. W. and Khorana, A. (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking & Finance*, 24(10), pp. 1651-1679. [https://doi.org/10.1016/S0378-4266\(99\)00096-5](https://doi.org/10.1016/S0378-4266(99)00096-5)
8. Chauhan, Y., Ahmad, N., Aggarwal, V. and Chandra, A. (2020). Herd behaviour and asset pricing in the Indian stock market. *IIMB Management Review*, 32(2), pp. 143-152. <https://doi.org/10.1016/j.iimb.2019.10.008>
9. Chiang, T. C. and Zheng, D. (2010). An empirical analysis of herd behavior in global stock markets. *Journal of Banking & Finance*, 34(8), pp. 1911-1921. <https://doi.org/10.1016/j.jbankfin.2009.12.014>
10. Choi, S. Y. (2021). Analysis of stock market efficiency during crisis periods in the US stock market: Differences between the global financial crisis and COVID-19 pandemic. *Physica A: Statistical Mechanics and Its Applications*, 574, A. 125988. <https://doi.org/10.1016/j.physa.2021.125988>
11. Christie, W. G. and Huang, R. D. (1995). Following the pied piper: do individual returns herd around the market? *Financial Analysts Journal*, 51(4), pp. 31-37. <https://doi.org/10.2469/faj.v51.n4.1918>
12. Demirer, R. and Kutan, A. M. (2006). Does herding behavior exist in Chinese stock markets?. *Journal of International Financial Markets, Institutions and Money*, 16(2), pp. 123-142. <https://doi.org/10.1016/j.intfin.2005.01.002>

13. Demirer, R., Leggio, K.B. and Lien, D. (2019) 'Herding and flash events: Evidence from the 2010 Flash Crash', *Finance Research Letters*, 31(1), pp. 1-22. <https://doi.org/10.1016/j.frl.2018.12.018>
14. Espinosa-Méndez, C. and Arias, J. (2021). COVID-19 effect on herding behaviour in European capital markets. *Finance Research Letters*, 38(1), A. 101787. <https://doi.org/10.1016/j.frl.2020.101787>
15. Espinosa-Méndez, C. and Maquieira, C. P. (2023). Herding behavior of business groups during COVID-19: Evidence in Chile. *Investigación Económica*, 82(324), pp. 98-125. <https://doi.org/10.22201/fe.01851667p.2023.324.83713>
16. Ferreruela, S. and Mallor, T. (2021). Herding in the bad times: The 2008 and COVID-19 crises. *The North American Journal of Economics and Finance*, 58(1), A. 101531. <https://doi.org/10.1016/j.najef.2021.101531>
17. Gao, X., Ren, Y. and Umar, M. (2022a). To what extent does COVID-19 drive stock market volatility? A comparison between the US and China. *Economic Research-Ekonomska Istraživanja*, 35(1), pp. 1686-1706. <https://doi.org/10.1080/1331677X.2021.1906730>
18. Gao, Y., Zhao, C., Sun, B. and Zhao, W. (2022b). Effects of investor sentiment on stock volatility: new evidence from multi-source data in China's green stock markets. *Financial Innovation*, 8(1), p. 77. <https://doi.org/10.1186/s40854-022-00381-2>
19. Ghorbel, A., Snene, Y. and Frikha, W. (2023). Does herding behavior explain the contagion of the COVID-19 crisis?. *Review of Behavioral Finance*, 15(6), pp. 889-915. <https://doi.org/10.1108/RBF-12-2021-0263>
20. Hudson, Y., Yan, M. and Zhang, D. (2020). Herd behaviour & investor sentiment: Evidence from UK mutual funds. *International Review of Financial Analysis*, 71, A. 101494. <https://doi.org/10.1016/j.irfa.2020.101494>
21. International Monetary Fund (IMF). (2023). World economic outlook: navigating global divergences. Washington, DC: International Monetary Fund. Available at: <https://doi.org/10.5089/9798400243431.081>
22. Javaira, Z. and Hassan, A. (2015). An examination of herding behavior in Pakistani stock market. *International Journal of Emerging Markets*, 10(3), pp. 474-490. <https://doi.org/10.1108/IJoEM-07-2011-0064>
23. Lee, Y. C., Wu, W. L. and Lee, C. K. (2021). How COVID-19 triggers our herding behavior? Risk perception, state anxiety, and trust. *Frontiers in Public Health*, 9, A. 587439. <https://doi.org/10.3389/fpubh.2021.587439>
24. Li, W., Chien, F., Waqas Kamran, H., Aldeehani, T. M., Sadiq, M., Nguyen, V. C. and Taghizadeh-Hesary, F. (2022). The nexus between COVID-19 fear and stock market volatility. *Economic research-Ekonomska istraživanja*, 35(1), pp. 1765-1785. <https://doi.org/10.1080/1331677X.2021.1914125>
25. Litimi, H., BenSaïda, A. and Bouraoui, O. (2016). Herding and excessive risk in the American stock market: A sectoral analysis. *Research in International Business and Finance*, 38, pp. 6-21. <https://doi.org/10.1016/j.ribaf.2016.03.008>
26. Mensi, W., Hammoudeh, S., Al-Jarrah, I. M. W., Sensoy, A. and Kang, S. H. (2017). Dynamic

- risk spillovers between stock markets. *North American Journal of Economics and Finance*, 41, pp. 133–155. <https://doi.org/10.1016/j.najef.2017.03.003>
27. Mishra, P. K. and Mishra, S. K. (2021). COVID-19 pandemic and stock market reaction: empirical insights from 15 Asian countries. *Transnational Corporations Review*, 13(2), pp. 139–155. <https://doi.org/10.1080/19186444.2021.1924536>
 28. Mishra, P. K. and Mishra, S. K. (2023). Do banking and financial services sectors show herding behaviour in Indian stock market amid COVID-19 pandemic? Insights from quantile regression approach. *Millennial Asia*, 14(1), pp. 54–84. <https://doi.org/10.1177/09763996211032356>
 29. Ramelli, S. and Wagner, A. F. (2020). Feverish stock price reactions to COVID-19. *The Review of Corporate Finance Studies*, 9(3), pp. 622–655. <https://doi.org/10.1093/rcfs/cfaa012>
 30. Sheikh, M. F., Bhutta, A. I. and Parveen, T. (2025). Herding or reverse herding: the reaction to change in investor sentiment in the Chinese and Pakistani markets. *International Journal of Emerging Markets*, 20(1), pp. 74–91. <https://doi.org/10.1108/IJOEM-02-2022-0270>
 31. Sihombing, N., Sadalia, I. and Wibowo, R. P. (2021). Analysis of herding behavior of investors in the banking industry in Indonesian stock market for February–June 2020 period. *International Journal of Research and Review*, 8(3), pp. 593–600.
 32. Tan, L., Chiang, T. C., Mason, J. R. and Nelling, E. (2008). Herding behavior in Chinese stock markets: An examination of A and B shares. *Pacific-Basin Finance Journal*, 16(1–2), pp. 61–77. <https://doi.org/10.1016/j.pacfin.2007.04.004>
 33. Vuong, G. T. H., Nguyen, M. H. and Keung Wong, W. (2022). CBOE volatility index (VIX) and corporate market leverage. *Cogent Economics & Finance*, 10(1), A. 2111798. <https://doi.org/10.1080/23322039.2022.2111798>
 34. Wang, Y., Xiang, Y., Lei, X. and Zhou, Y. (2022). Volatility analysis based on GARCH-type models: Evidence from the Chinese stock market. *Economic research-Ekonomska istraživanja*, 35(1), pp. 2530–2554. <https://doi.org/10.1080/1331677X.2021.1967771>
 35. Wang, Y., Yu, C. and Zhao, X. (2023). Does herding effect help forecast market volatility? Evidence from the Chinese stock market. *Journal of Forecasting*, 42(5), pp. 1275–1290. <https://doi.org/10.1002/for.2968>
 36. Welch, I. (1992). Sequential sales, learning, and cascades. *Journal of Finance*, 47(2), pp. 695–732. <https://doi.org/10.1111/j.1540-6261.1992.tb04406.x>
 37. Wu, G., Yang, B. and Zhao, N. (2020). Herding behavior in Chinese stock markets during COVID-19. *Emerging Markets Finance and Trade*, 56(15), pp. 3578–3587. <https://doi.org/10.1080/1540496X.2020.1855138>
 38. Yuan, H. (2021). *Herding effect in China's A-share stock market under COVID-19 pandemic*. In *Proceedings of the 2021 7th International Conference on E-Business and Applications (ICEBA 2021)* (pp. 70–76). Association for Computing Machinery (ACM), New York. <https://doi.org/10.1145/3457640.3457662>
 39. Zarembo, A., Kizys, R., Aharon, D. Y. and Demir, E. (2020). Infected markets: Novel coronavirus, government interventions, and stock return volatility around the globe. *Finance Research Letters*, 35, A. 101597. <https://doi.org/10.1016/j.frl.2020.101597>
 40. Zhang, H. and Giouvris, E. (2022). Measures of volatility, crises, sentiment and the role of US

‘Fear’Index (VIX) on herding in BRICS (2007–2021). *Journal of Risk and Financial Management*, 15(3), p. 134. <https://doi.org/10.3390/jrfm15030134>



RESEARCH ARTICLE

Ranking the Factors Affecting the Level of Tax Compliance within the Framework of Comprehensive Tax Plan

Rahman Sajadpour, Mehdi ArabSalehi*, Iraj Kazemi

Department of Administrative Sciences and Economics, University of Isfahan, Isfahan, Iran

How to cite this article:

Sajadpour, R. , Arab Salehi, M. and Kazemi, I. (2026). Ranking the Factors Affecting the Level of Tax Compliance within the Framework of Comprehensive Tax Plan. Iranian Journal of Accounting, Auditing and Finance, 10(3), 21-47. doi: 10.22067/ijaaf.2026.48222.1611
https://ijaaf.um.ac.ir/article_48222.html

ARTICLE INFO

Article History

Received: 2025-11-28

Accepted: 2026-01-22

Published online: 2026-07-12

Abstract

According to this study, one of the most significant information technology projects of Iran's National Tax Administration (INTA) is the comprehensive tax plan, which considers the factors influencing tax compliance. In this research, the factors affecting tax compliance, including accounting variables (internal factors) and environmental conditions (external factors), were evaluated and ranked using decision tree algorithms. This study employed a longitudinal firm-year dataset (panel data) and analyzed it using machine-learning techniques rather than panel econometric models. The comparison of decision tree models revealed that the variables influencing tax compliance, in descending order of importance, were disclosure quality, board remuneration, the percentage of free-float shares, government ownership and influence in firms, cash-holding level, income smoothing, and firm size. The ranking of influential variables in this study provided valuable guidance for regulatory and supervisory institutions seeking to enhance tax compliance. Moreover, the findings indicated an evaluative indicator of the implementation level of the comprehensive tax plan, which can be used to assess the performance of the Iranian National Tax Administration and for future research.

Keywords:

Disclosure Quality, Financial Reporting Transparency, Tax burden; Tax compliance, Tax Plan

<https://doi.org/10.22067/ijaaf.2026.48222.1611>



NUMBER OF REFERENCES
36



NUMBER OF FIGURES
2



NUMBER OF TABLES
11

Homepage: <https://ijaaf.um.ac.ir>
 E-ISSN: 2717-4131
 P-ISSN: 2588-6142

*Corresponding Author: Mehdi ArabSalehi
 Email: mehdi_arabsalehi@ase.ui.ac.ir
 Tel: 09133289157
 ORCID: 0000-0003-3982-9185

1. Introduction

One of the main ways the government makes money is through taxes. This is a strategic, fundamental, and evolving subject that receives significant attention to achieve development goals, boost economic growth, ensure fair income distribution, expand public services, and stabilize social welfare. In a situation of falling oil revenues, the country depends on it; therefore, it is essential to maximize tax collection. In addition, higher-level national policy documents, particularly the economic transformation plan, Five-Year development programs, and the Twenty-year vision document, have emphasized that more tax revenue should be raised to reduce the share of oil income in government current expenditure (INTA, 2014). One of the essential tax indicators is taxpayer compliance, which measures people's willingness to pay taxes or evade them. Based on the national comprehensive tax plan, it is possible to determine and prioritize tax compliance based on collection amount, collection time, and collection expenses. The primary gain from paying one's fair share of taxes is the additional revenue that accrues to the government.

Higher tax compliance improves the efficiency of the tax system by increasing voluntary tax payments, enhancing tax revenues relative to GDP and budget targets, and reducing tax collection costs. As a result of higher compliance, taxes are collected more quickly, since they are less likely to be disputed and enforced, which is particularly important in inflationary situations (Kirchler et al., 2008; OECD, 2022). Among the three important tax indicators is taxpayer compliance, which indicates whether people are willing and able to pay their taxes. Moreover, higher compliance mitigates tax evasion and strengthens perceptions of fairness within the tax system, thereby preventing negative spillover effects among taxpayers (Torgler, 2007; INTA, 2014).

In its pursuit of a more contemporary tax system, the Iranian National Tax Administration has embarked on several large-scale information technology projects, including the comprehensive tax plan. Taxpayer operations, including registration, return filing, tax payment, auditing, and tax calculation, are performed entirely digitally. Proponents of this idea also claim that it will lead to less tax evasion, more tax compliance, and better tax equality, all of which will make taxpayers happier.

According to the plan's designers, three primary objectives have been envisioned for the comprehensive tax plan. To increase tax revenues, the first goal is to prevent tax evasion, improve tax compliance, and eliminate unnecessary tax exemptions (Allingham and Sandmo, 1972; Alm and Torgler, 2011). The second goal is to lower operational expenses by making processes more efficient and automating them in the Iranian National Tax Administration. The third goal is to make it easier for stakeholders to work with the organization, which will help make tax fairness more likely and increase taxpayer respect and happiness (Torgler, 2007; INTA, 2014).

One of the main reasons for making the comprehensive tax plan was that taxpayers weren't following the rules and paying their taxes. This shows that tax compliance remains a major concern for both the government and the Iranian National Tax Administration.

There are 32 projects in the comprehensive tax plan, which are divided into five main areas: the integrated tax system, infrastructure, human resources, stakeholders, and project management. Since 2009, this strategy has been gradually implemented, and the country's tax system has undergone numerous changes.

The purpose of this study is to seek an answer to the following question: What motivates individuals to pay their taxes without protest? Several factors have been linked to taxpayers' propensity to pay their fair share of taxes. These include taxpayers' faith in the government, the severity of corruption, the taxpayers' financial situation, the openness of financial reporting, the relationship between tax rates, disclosure of tax information, and corporate governance practices, taxpayers' prior interactions with public services, taxpayers' religious beliefs, the degree to which the tax system is integrated, and many more. The magnitudes of these characteristics differ, though.

Financial openness, disclosure quality, firm size, cash holdings, ownership structure, income smoothing, and other firm-level characteristics are considered internal factors in this study. Managerial decisions and internal corporate policies largely determine these factors.

Institutional and environmental variables that are not internal to individual businesses but are influenced by the larger regulatory and tax climate are known as external factors. Examples of such variables include the total tax burden, the extent to which the Comprehensive Tax Plan has been implemented, and the degree to which individual firms have been impacted. Previous research on tax compliance in Iran has mostly examined its causes in isolation rather than within the institutional framework of the Comprehensive Tax Plan, despite a growing body of literature on the topic. To be more specific, no study has yet used an integrated analytical model to rank the factors that influence tax compliance or to separate internal firm-level elements from external macroeconomic and institutional ones. There has also been a lack of quantitative measurement and the inclusion of the Comprehensive Tax Plan's implementation timeline and extent as an explanatory variable in previous empirical research. This study addresses these gaps by developing a comprehensive framework that integrates internal and external determinants of tax compliance alongside a novel, data-driven measure of the implementation of the Comprehensive Tax Plan, thereby providing a multidimensional and context-specific analysis of tax compliance in Iran. This study contributes to the literature in several ways. From a scientific perspective, the study provides one of the first integrated empirical frameworks that combines internal corporate governance and financial variables with an externally constructed, time-varying indicator of the Comprehensive Tax Plan. Using data mining and decision-tree techniques, the study uncovers nonlinear correlations and variable ranks that are difficult to capture with standard methods, thereby going beyond average impact estimates. The results support more targeted and efficient enforcement and policy design by providing practical insights for tax authorities and policymakers and identifying which firm-level and institutional factors are most informative about tax compliance behavior. The remainder of the article is structured as follows. Section 2 reviews the theoretical foundations and related empirical literature on tax compliance. Section 3 describes the research design, data sources, variable construction, and modeling framework. Section 4 presents the empirical results and compares the performance of the applied models. Section 5 concludes the study by discussing the results, outlining the limitations, and offering some final thoughts and suggestions for further study. Panel data collected from businesses between 2009 and 2022 are used for the empirical analysis.

2. Theoretical foundations

The Iranian National Tax Administration, along with the parliament and the government, has been working on policies, regulations, and digital systems to address taxpayer noncompliance and the high cost of tax collection. These efforts aim to increase overall tax revenues, decrease the cost of tax evasion, and improve taxpayer compliance (INTA, 2014). In particular, due to economic sanctions, revenues from sources such as oil and petroleum product sales have declined substantially, thereby making reliance on taxation increasingly critical.

To improve tax compliance, it is necessary to identify and understand the factors that influence taxpayers' attitudes and tendencies towards paying taxes. Tax compliance is influenced by several determinants, as discussed in the following sections.

2.1. The economic theory of crime

The comprehensive tax plan is one of the most important factors affecting people's willingness to

pay their taxes. It is argued that people have to choose between getting caught and paying the costs of being caught, or not getting caught and getting the advantages of not getting caught. Portfolio theory suggests that in the "tax evasion game," smart people aim to achieve the best possible return. So, the potential financial benefits of breaking the rules or not paying taxes are weighed against the potential losses from getting caught and punished. People pay their taxes because they are terrified of getting caught and penalized, according to this way of thinking. In other words, how many audits there are and how likely it is that someone will "win" or not get caught affect how many people pay their taxes. The amount of the fine if caught also matters ([Allingham and Sandmo, 1972](#)). The model assumes that a person has a steady income and has to choose how much of it to report to the government (R) and how much to keep secret. The person pays a proportional tax rate t on each recorded unit of income. There is no tax on unreported income, but there is a steady chance (p) that the person will be audited. If that happens, they will have to pay a fine (f) for each unit of unreported income, in addition to the tax they owe.

The individual's net income, if underreported income is detected, is given by *Equation (1)*.

$$\text{Equation (1). } IC = I - tR - f[t(I - R)]$$

The individual's net income when unreported income is not detected is given by the following equation: $IN = I - tR$. It is assumed that the individual chooses the level of reported income (R) to maximize expected utility. Accordingly, the individual's optimal expected income function is expressed as *Equation (2)*.

Where I is actual income, p shows the probability of detection, f denotes the penalty rate, R presents reported income, and t indicates the tax rate.

$$\text{Equation (2). } EU(I) = pU(I_C) + (1 - p)U(I_N)$$

The economic theory of crime postulates that the likelihood of detection (p) and the severity of punishments (f) are the primary determinants of tax compliance. Because of its emphasis on digitization and information system integration, the Comprehensive Tax Plan has an operational impact on both criteria. The economic theory of crime postulates that the likelihood of detection (p) and the severity of punishments (f) are the primary determinants of tax compliance. Because of its emphasis on digitization and information system integration, the Comprehensive Tax Plan has an operational impact on both criteria. In addition, real-time data access and systematic documentation improve the enforceability of tax regulations, thereby raising the expected cost of non-compliance by strengthening the credibility and application of penalties (f). Consequently, the Comprehensive Tax Plan functions as an institutional mechanism that directly links digital infrastructure to the core determinants of tax compliance in the [Allingham and Sandmo \(1972\)](#) framework. The economic theory of crime, through comparative statistical analysis, demonstrates that reported income (compliance) increases with higher probabilities of detection or higher penalty rates. This approach further concludes that individuals pay taxes primarily due to the economic consequences of detection and punishment.

A wide range of instruments and approaches has been developed to enhance the likelihood of detecting tax evasion. One of the most important developments in this area has been the use of technology in tax administration, implemented gradually through several programs.

This process has been going on in Iran since 2009 as part of the comprehensive tax plan.

2.2. The slippery slope model

The tax rate is another subject that affects how well people pay their taxes. Figure 1 shows that the slippery slope model of tax compliance uses both economic reasoning and human psychology to explain why people choose to follow tax rules, either willingly or out of necessity. The approach is based on two basic ideas: how much people trust the government and how strong they think the

government is. People's ideas about how much power the government has are shaped by things like how high taxes are, how likely they are to be audited, and how rigorous the punishments are. [Kirchler et al. \(2008\)](#) found that psychological factors like taxpayers' perception of fairness affect the amount of trust they have in authority ([Kirchler et al., 2008](#)). In the slippery slope paradigm, taxpayers may view an increase in the tax rate as unjust, thereby reducing tax compliance. However, when confidence is high, such gains are seen as community consensus ([Kirchler et al., 2008](#)). Tax compliance behavior is not directly tested in the empirical model; instead, the economic and institutional mechanisms underlying it are captured through observable proxies.

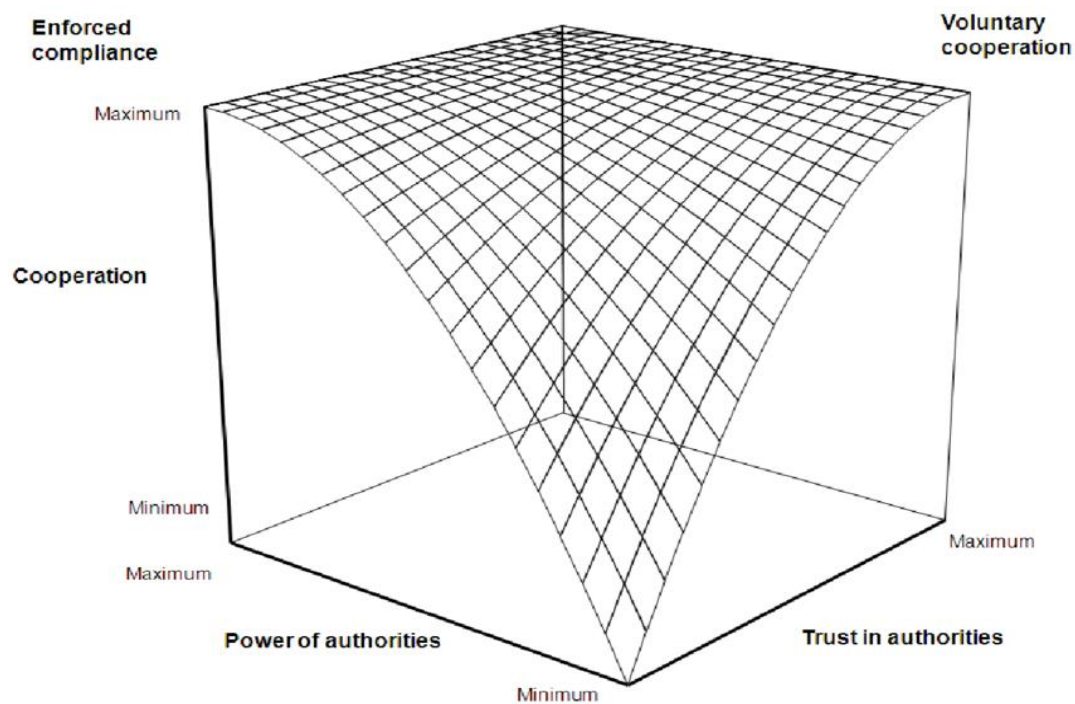


Figure 1. Relationship between the level of compliance (voluntary and enforced) and trust or power of authorities ([Kirchler et al., 2008](#))

2.3. The tax burden

The tax burden is an economic constraint imposed on business entities through tax deductions from their budgets. This implies that taxation takes a particular portion of income that could otherwise be allocated to development activities by corporations or individual entrepreneurs. The concept of the tax burden should therefore be defined at both the macroeconomic and microeconomic levels. At the firm level, the tax burden is generally divided into two categories: tax on profit and tax on value added. As the tax burden increases, tax compliance among businesses is expected to decrease, as they seek to preserve conditions conducive to growth and entrepreneurship. In this study, since other research variables are already related to profit taxation and following [Block \(2016\)](#), the value-added tax rate and charges are employed as measures of the tax burden.

2.4. Corporate governance

Corporate governance is another internal factor that affects tax compliance, and corporate policy can be considered as a range. The narrow approach to corporate governance looks at the firm's relationship with its owners. The broad approach, on the other hand, looks at everyone's rights.

Companies have responsibilities to more than just the people who currently have a stake in them. They also have these responsibilities to future generations, society as a whole, and the environment. The broader strategy is consistent with stakeholder theory, whereas the narrower strategy is consistent with agency theory. Among the many major players in the limited approach is the government. Thus, before distributing any further profit, a company should determine and set aside the portion that goes to the government, which is primarily derived from taxes (Slemrod, 2019; Kerr et al., 2016; Duhoon and Singh, 2023). Any company that does not want to be responsible for other people's rights is still legally required to respect the government's rights by determining and paying the correct amount of tax. From a narrow perspective, the interests of those who prepare information may not align with those of other parties, such as the government, which stands to gain the most, due to a lack of transparency and weak corporate governance. In these situations, businesses see their responsibility only as staying out of trouble with the law and paying their taxes. On the other hand, more general meanings of corporate governance stress that the company has many connections with many different groups of people who have a stake in it. From this point of view, corporations are responsible for more than just their shareholders. They are also responsible for other groups affected by their actions. This method is based on principles such as openness, fairness, and the protection of stakeholders' rights. Accountability is considered an essential part of corporate governance in this way of thinking. Firms should follow their social responsibility by paying their taxes, as part of their overall duty to society (Abdelfattah and Aboud, 2020).

2.5. Quality of disclosure

The quality of disclosure and financial reporting also affects tax compliance. Agency theory suggests that tax decisions may be made in managers' best interests when ownership and management are independent. Thus, reducing tax avoidance-related agency costs is crucial for owners. According to managerial opportunism (Balakrishnan et al., 2019), managers evade taxes to hide their opportunistic actions. Improving transparency can save agency costs and fill information gaps. An organization's transparency is proportional to the quality of its financial reporting and disclosure policies. Company managers and outside investors are already at a disadvantage due to the lack of transparency that results from aggressive tax methods, which further compounds the problem. Tax avoidance activities could provide managers with tools for dishonest behavior, such as changing wages, doing business with related parties, and diverting resources (Zhang et al., 2025). While tax avoidance strategies may save the company money on taxes, they also complicate the company's finances and organization. Often, this makes it harder to communicate with external parties such as investors, creditors, and analysts, thereby making financial reporting less clear. That is, less transparency makes it worse because investors, shareholders, creditors, and managers do not have the same amount of knowledge. One possible cost of tax planning is that it can make a company's financial reports less clear (Balakrishnan et al., 2019). Balakrishnan et al. (2019) stated that firms have to pay a high price when there is insufficient transparency. These costs include higher costs of capital (both stock and debt), worsening governance problems, and less efficient investment. Because of this, the better the exposure, the less tax avoidance there is, which causes problems for the government.

2.6. Cash holdings

The taxpayers' ability to pay their taxes is another internal factor that affects their obedience. If managers have no conflicts of interest, they seek to increase the company's value and maintain an appropriate level of cash on hand for two main reasons: the transaction motive and the precautionary motive (Opler et al., 2001). The transaction motive holds that companies that keep cash on hand can

meet their current obligations without borrowing from outside sources or selling assets, which can be expensive. The precautionary motive, on the other hand, holds that companies should keep cash on hand to take advantage of investment opportunities that could generate returns. The opportunity costs of holding cash, such as the lost returns from investing in long-term investments, are weighed against these benefits. According to this theory, there should not be a clear link between a company's decision to avoid paying taxes and its cash holdings, as long as the optimal level of cash is determined by comparing the marginal benefits and costs of liquidity. When ownership and management are kept separate, managers may make choices that are not in the best interests of shareholders. Cash and other easily spent assets can be easily abused by managers who only care about themselves. Therefore, ownership disputes can accelerate the spending of a company's cash assets. Thus, it stands to reason that managers might avoid paying taxes to keep cash on hand within the company (Foley et al., 2007). In turn, this behavior makes them less likely to fulfill their tax obligations, since doing so would reduce the cash they have available.

2.7. Theoretical background and empirical implications

In this study, an integrated framework that connects the empirical decision tree and ranking analysis to the factors determining tax compliance is constructed by synthesizing multiple theoretical viewpoints. Allingham and Sandmo's (1972) economic model of crime serves as a basis for incorporating enforcement-related elements into the Comprehensive Tax Plan's digitalization, thereby raising the likelihood of detection and the anticipated cost of non-compliance. Second, the slippery slope paradigm emphasizes institutional credibility and transparency, highlighting the importance of trust and perceived legitimacy. The tax burden theory explains how higher effective tax pressure may weaken voluntary compliance, justifying the inclusion of tax burden indicators as external determinants. Fourth, corporate governance theory emphasizes that stronger governance mechanisms and ownership structures hinder opportunistic behavior, leading to higher compliance. Fifth, transparency and disclosure theories suggest that improved information quality reduces information asymmetry and aggressive tax behavior. Finally, the cash holdings theory posits that firms with greater liquidity may face different compliance incentives due to increased monitoring capacity and opportunity costs. Together, these theoretical blocks motivate the selection of internal and external variables and form the basis for the decision tree analysis, which ranks the relative importance of determinants based on their explanatory power for tax compliance.

2.7.1. Empirical background

2.7.1.1. Behavioral & psychological factors

This study examines noncompliance with business income tax and ranks the factors influencing it. The report is published by Kaghazloo and Borrego (2022). The Analytic Hierarchy Process (AHP) method was used to analyze data gathered via a grid questionnaire using the Saaty pairwise comparison technique. The findings showed that knowledge inequality, behavior based on reference groups, and third-party reporting were the most important factors leading to tax evasion.

Bani-Khalid et al. (2022) investigated the determinants of tax compliance among small and medium-sized enterprises (SMEs) in Jordan, focusing on the Theory of Planned Behavior (TPB) as the theoretical framework. An analysis of 385 valid responses was obtained from 660 questionnaires distributed to managers and owners of manufacturing SMEs in Jordan using a systematic random sampling method. To evaluate the measurement and structural models and assess the hypothesized correlations within the proposed framework, the research utilized Partial Least Squares Structural Equation Modeling (PLS-SEM). The findings indicated that attitude toward behavior, subjective norms, perceived behavioral control, and patriotism were significant determinants of the intention to

comply with tax obligations among the surveyed firms. While these studies primarily emphasize behavioral and perceptual factors, the present study captures their empirical implications using observable firm-level and institutional proxies such as governance, transparency, tax burden, and enforcement mechanisms thereby enabling these influences to be systematically ranked within a decision-tree framework.

Taing and Chang (2021), in their study titled “Determinants of tax compliance: a focus on the theory of planned behavior,” applied the theory of planned behavior (TPB) to examine citizens’ intention to comply with tax obligations in Phnom Penh, Cambodia. The TPB posits that three key components underlie behavioral intention: attitude, subjective norms, and perceived behavioral control. Based on these constructs, seven determinants were selected for the study: tax morale, tax fairness and trust in government, perceived power, tax complexity, tax information, and tax awareness. To investigate these determinants, the authors collected survey data from 402 respondents in 2016. The results of this study supported the view that tax morale, tax fairness, and tax complexity have significant effects on citizens’ intention to comply with tax obligations. On the other hand, authority power, trust in government, and tax information and awareness were not statistically significantly related.

Mebratu (2024) examined the theoretical basis of studies on tax compliance in some developing countries. This study examined prior research on tax compliance in developing countries worldwide from 1998 to 2022, both by topic and by time period. The goal was to identify the most widely studied factors affecting tax compliance. The study stressed that long-term tax compliance can only be achieved when fiscal policy takes into account the key factors that influence people’s decisions to comply. One of the biggest problems in tax management, especially in developing countries, is getting people to pay their taxes. The review found that taxpayers’ attitudes and perceptions of tax fairness were each identified six times as major determinants of compliance. In addition, tax information, how people are expected to act, and subjective norms were each mentioned four times. Three times, income sources, tax law complexity, moral sensitivity, and religious understanding were named as important factors. Two times, perceived tax system structure and perceived usefulness were named as factors.

Alm and Torgler (2011) investigated the role of ethics and cultural values in shaping tax compliance behavior. Using comparative data from different countries, they showed that moral and normative factors affect taxpayer behavior as strongly as, or more strongly than, economic factors. A high rate of tax compliance is indicative of a society with strong ethical standards and a high level of social trust. If people do not trust the government or believe many institutions are corrupt, even when there are strict tax laws, they will comply with them less (Alm and Torgler, 2011).

2.7.1.2. Trust, Fairness & Tax Morale

Bartha and Boda (2024) conducted an international comparison and studied the effects of corruption scandals and political legitimacy crises on tax morale in Hungary. Results suggest that if public trust is substantially undermined, tougher enforcement methods, such as greater audits or penalties, may not effectively improve compliance. The paper also distinguishes between compliant behavior, which is enforced, deriving from fear of sanctions, coping with a lack of trust and perceived fairness, vis-à-vis compliance essentially being voluntary; when corruption exists, then voluntary compliance tends towards zero (Bartha and Boda, 2024). These findings justify the inclusion of external institutional variables related to transparency, enforcement intensity, and the probability of detection in the present study’s ranking of tax compliance determinants.

Mohammed and Tangi (2023) employed an analytical approach and identified major determinants of tax compliance, with attention to tax justice, transparency, and the quality of government–taxpayer

relations. Results indicate that tax legislation and administrative procedures are overly complex, leading to decreased compliance. Additionally, collectivist attitudes toward the tax burden and taxes, in the form of a weight sense among taxpayers, also influence willingness to comply. This study suggests that policymakers should use a multi-faceted strategy for enforcement, including punitive measures and tax education, to increase institutional credibility ([Mohammed and Tangi, 2023](#)). This evidence supports the inclusion of tax burden, transparency, and institutional quality as key external determinants in the empirical model and motivates their joint evaluation in the ranking analysis.

2.7.1.3 Enforcement, Audits & Policy Interventions

[Nwokoye et al. \(2023\)](#) examined how financial rewards affect the way Nigerian industrial firms handle taxes. A structured questionnaire and a multistage sampling method were used to get information from 800 companies in three industry clusters. The poll answers were analyzed using descriptive statistics and logistic regression. Regular tax audits, firm size, making tax-related communications easier to understand, delivering deterrent messages, entrepreneurs' education, and the legitimacy of the current government all affect enterprises' tax compliance. Financial incentives such as tax credits, tax reductions, capital allowances, and investment incentives also play a significant role. These findings directly motivate the inclusion of firm size as an internal determinant and audit probability and government-related factors as external determinants in the ranking framework.

[De Neve et al. \(2021\)](#) analyzed data from large-scale studies conducted in Belgium to determine the effects of tax simplification, nonpayment prevention, and tax morale on taxpayer compliance. The researchers altered the tax office's communication with Belgian income taxpayers by conducting four field experiments. The results demonstrated that compliance rates were substantially higher when communications were made more understandable, and the effect was even more pronounced when the messages were intended to discourage people. However, bringing up tax attitudes seldom helped and often made matters worse. The study also showed that steps to simplify processes are much more cost-effective, which means significant savings in administrative costs.

[Slemrod \(2019\)](#) provided a comprehensive review of the literature on tax compliance and related enforcement policies. The study surveyed hundreds of empirical and theoretical works, which proposed an analytical framework to understand the interactions among enforcement policies (audits, penalties, and incentives), taxpayer characteristics (firm size, liquidity, and ownership structure), and the institutional context. Based on the assessment findings, it is clear that a mix of "hard" enforcement instruments and "soft" techniques, such as trust-building and procedural simplicity, is necessary for an effective strategy to increase compliance. The paper also highlighted disparities between rich and poor nations, demonstrating how local institutional factors determine the success of tax policies ([Slemrod, 2019](#)).

2.7.1.4. Corporate Governance & Firm-Level Determinants

[Desai and Dharmapala \(2006\)](#) show that corporate governance also has an indirect effect on tax avoidance through the integrity of financial reporting, which serves as a mediating factor. Furthermore, higher audit quality was found to reduce tax avoidance, both directly and indirectly, via its positive influence on financial reporting quality.

Other studies have shown that board independence, executive incentives, tax penalties, tax position, tax expertise of managers, managerial risk aversion, firm performance, firm size, and the difference between reported taxable income and accounting income all affect taxpayer compliance or avoidance. These studies include [Armstrong et al. \(2015\)](#) and [Alasfour et al. \(2016\)](#).

frameworks, including behavioral models such as the Theory of Planned Behavior, which

emphasizes attitudes, norms, and perceived behavioral control; the psychological tax contract, which highlights trust and reciprocity between taxpayers and tax authorities; and the slippery slope framework, which integrates trust with enforcement power. A comparative reading of this literature reveals consistent patterns across studies: trust, fairness, and transparency are central drivers of voluntary tax compliance. At the same time, governance-related mechanisms—such as board characteristics and ownership structure constrain opportunistic behavior and aggressive tax practices. At the same time, enforcement-related factors, including tax burden and audit probability, exhibit context-dependent and sometimes nonlinear effects, as stronger enforcement may enhance compliance in some settings but weaken voluntary compliance under excessive tax pressure. Despite these insights, prior studies have largely examined compliance determinants in isolation and outside the operational framework of Iran's Comprehensive Tax Plan and digital tax system, without systematically comparing internal firm-level factors with external institutional determinants or ranking their relative importance. Addressing these gaps, the present study adopts a firm-level and institutional perspective by embedding tax compliance analysis within the framework of the Comprehensive Tax Plan and its digital infrastructure, and by applying a decision tree-based approach to jointly evaluate and rank observable internal and external determinants such as governance, transparency, liquidity, ownership structure, tax burden, and audit probability thereby capturing nonlinear interactions and offering a structured prioritization of compliance drivers not provided by earlier Iranian studies.

While the literature on tax compliance extensively emphasizes behavioral constructs such as trust in government, perceived fairness, and corruption, these variables are inherently latent and not directly observable in firm-level archival data. Accordingly, this study adopts a proxy-based empirical strategy, whereby observable firm-level and institutional variables are used to capture the economic and informational manifestations of these underlying behavioral mechanisms. Variables such as disclosure quality, corporate governance attributes, and government ownership reflect transparency, monitoring, and institutional credibility, while the ICTPlan index captures the evolution of enforcement capacity and system digitalization. This approach enables the integration of behavioral theory within a data-driven empirical framework.

2.8. Critique of previous studies and research innovation

Many prior studies (e.g., [Mebratu, 2024](#); [Bani-Khalid et al., 2022](#); [Nwokoye et al., 2023](#); [Bartha and Boda, 2024](#)) have focused either on a limited set of economic or accounting variables or on behavioral and psychological factors, without integrating these dimensions with contextual environmental and institutional variables related to the Comprehensive Tax Plan. In Iran, researchers have similarly relied on a narrow set of economic or accounting indicators, often using financial ratios and corporate governance measures comparable to those employed in global studies of fraudulent financial reporting (e.g., [Armstrong et al., 2015](#); [Alasfour et al., 2016](#)). As a result, two important gaps remain in the literature. First, prior studies have not sufficiently incorporated economic, environmental, and institutional factors, particularly those associated with the ongoing implementation of the Comprehensive Tax Plan, alongside internal accounting variables in tax compliance prediction models. Second, existing research has largely relied on traditional statistical approaches and has not employed data-driven ranking methods to assess the relative importance of internal and external determinants. Addressing tax compliance in Iran, therefore, requires a comprehensive framework that jointly considers internal firm-level characteristics and external environmental factors and systematically ranks their contributions.

3. Research methodology

The goal of this study is to find and rank the factors that affect tax compliance for businesses listed on the Tehran Stock Exchange under the comprehensive tax plan. The study uses both accounting and non-accounting data and will consider internal and external factors. The study population consists of all companies listed on the Tehran Stock Exchange. This study employs a longitudinal firm–year dataset (panel data) and analyzes it using machine-learning techniques rather than panel econometric models. The final dataset comprises 1,337 firm–year observations. The sample of listed firms was selected using a systematic elimination approach. Firms were included if they had a fiscal year ending in March, did not change their fiscal year during the study period, had the required financial information available, and were listed on the Tehran Stock Exchange no later than the end of fiscal year 2009. Financial institutions, including banks, insurance companies, investment firms, financial intermediaries, holding companies, and leasing firms, were excluded due to differences in corporate governance structures and financial disclosure requirements. Data were collected from the CODAL system, audited financial statements, and official records of the Securities and Exchange Organization. To extract information from the data, examine what drives corporations to comply with tax rules and pay taxes, and develop ways to make this happen nationwide, the study used a three-step data mining process. The first step was to understand and prepare the data. At this step, we examined values outside the normal range, codes that did not make sense for qualitative variables, outliers, missing data, data integration, and descriptive statistics. During the modeling stage, decision tree algorithms were used to examine and rank the factors that affect tax compliance levels. A hold-out validation approach was used to test the model, with 70% of the data being used for training and 30% for testing. The dependent variable (tax compliance) exhibits an imbalanced class distribution, with the compliant class occurring approximately 3 times as often as the noncompliant class. To address this issue, multiple strategies were employed. First, model performance was evaluated using metrics robust to class imbalance, including recall, precision, F1-score, AUC, and lift, rather than relying solely on accuracy. Second, cost-sensitive learning was applied in decision tree–based models by assigning a higher misclassification cost to the minority class, proportional to the observed class imbalance.

In this study, the factors influencing tax compliance, comprising both accounting and non-accounting data, were analyzed by considering internal and external variables. These variables are tax compliance, implementation of performance evaluation of the comprehensive tax plan, proportion of non-directional board member, amount paid to the board members as salary and wage, separation of CEO and chairman positions, public percentage ownership (free float), institutional accommodation ratio, government's presence on company and effects it has on companies decision making investment size firm size we consider Cash holding level in such a way that is more than 0.36, tax surcharge again from 05-02-2006 after passing law against this constitution type disclosure quality smoothing earnings Smoothed corporate income audit report form me suitable Operating Cash Flow difference result estimated operating cash flows discretionary accruals Total Accruals evaluation performance Evaluation Total Accruals from Differences ImplementationA comprehensive plan to meet the Islamic Republic for Iran. In conclusion, IBM SPSS Modeler Software was used to rank decision tree models on data from firms listed on the Tehran Stock Exchange. To construct the dependent variable, we first compute tax evasion for firm i in year t as the difference between the final tax assessed by the tax authority and the tax declared by the firm. To account for systematic differences across firms due to accounting standards and tax regulations, the sample mean of tax evasion is subtracted from each observation. Negative values, which indicate full compliance or overpayment, are set to zero, while positive values are retained. The resulting values

are then normalized to the $[0,1]$ interval to obtain a continuous non-compliance index (TA_{it}). Finally, tax compliance is defined as the inverse of this index:

$$Tax\ Compliance_{it} = 1 - TA_{it}$$

For classification in the decision tree model, a binary variable is constructed such that firms with $TA_{it} = 0$ are classified as compliant (1), and firms with $TA_{it} > 0$ are classified as non-compliant (0) (Dyreng et al., 2008).

The ICTPlan variable is constructed as a composite index based on expert assessments of the Comprehensive Tax Plan's implementation. The index was developed through a structured multi-step process. First, data were collected from a questionnaire covering 32 core projects across five main dimensions of the plan. For each project, experts evaluated (i) the year of effective implementation and (ii) the intensity of its impact on improving the tax system using a five-point Likert scale. The questionnaire was completed by 35 tax experts, including senior administrators and specialists in tax policy.

A panel of experts confirmed the instrument's content validity, and its reliability was assessed using Cronbach's alpha ($\alpha = 0.958$). Next, responses were numerically coded, and average impact scores were computed for each project. These scores were then aggregated at the dimension level to capture differences in implementation across the plan's five core areas. Finally, dimension-level scores were combined to construct an annual composite index, which was normalized to a 0–100 scale. This index provides a time-varying measure of the degree of implementation of the Comprehensive Tax Plan and is used in this study as an external institutional variable. A full methodological description of the index construction is provided in a companion paper by the same authors, currently under review.

Although the dataset has a panel (firm–year) structure, the objective of this study is not to estimate causal relationships using panel econometric techniques, but to identify and rank determinants of tax compliance using data-driven classification algorithms. Therefore, decision tree methods are employed, which do not require the assumptions or diagnostic tests associated with panel regression models.

The classification algorithm was trained and verified using a confusion matrix in this work. To assess the model, performance metrics were calculated from the confusion matrix, including accuracy (Acc), error rate (1-Acc), true positive rate (Recall⁺), positive predictive value (Precision⁺), negative predictive value (Precision⁻), and negative predictive value (Precision⁻). Furthermore, the F1 score and the Lift Index were used to evaluate the models' overall performance (Sokolova and Lapalme, 2009).

4. Research findings

4.1. Descriptive statistics

Table 1 displays the descriptive statistics of the research variables. These statistics include the mean, standard deviation, skewness coefficient, median, mode, number of observations, the lowest and highest values, and the research variables. Given that most variables in the study are continuous and multidimensional, outlier treatment was applied prior to discretization using a multivariate anomaly-detection approach. Specifically, a clustering-based technique was employed to partition the data into homogeneous peer groups, maximizing within-cluster similarity and between-cluster dissimilarity. For each observation, the distance from the center of its assigned cluster (anomaly index) was computed and evaluated relative to its peer group. Observations exhibiting unusually large distances from their cluster centers were identified as outliers and removed from the dataset. Regarding missing data, only one missing value was detected in the CashHold variable, which was replaced using the median value to preserve sample size and robustness.

Table 1. Descriptive statistics of the study variables

Symbol	Variable	Min	Max	Mean	SD	Skewness	Median	Mode	Uniqueness	Observations
OMBS	Non-executive board members	0.000	1.000	0.568	0.283	-0.804	0.600	0.600	--	1337
BOBO	Board remuneration	0.000	0.079	0.001	0.003	18.287	0.000	0.000	--	1337
INDCH	CEO-board separation	0	1	--	--	--	--	0	2	1337
FRFL	Free float share percentage	0.910	97.820	25.492	18.245	1.343	21.600	1.000	--	1337
POIN	Institutional ownership percentage	0.000	99.090	74.219	18.365	-1.351	78.190	99.000	--	1337
SOIC	Government ownership and influence	0.000	98.040	6.522	14.833	2.987	0.000	0.000	--	1337
Size	Firm size	9.069	20.467	14.253	1.616	0.595	14.076	12.212	--	1337
CashHold	Cash holding level	0.000	2.669	0.085	0.141	6.829	0.042	0.003	--	1337
Disclos	Disclosure quality	0.837	99.859	69.530	19.160	-0.676	73.471	50.000	--	1337
SMO	Earnings smoothing	0.027	34.721	1.659	2.875	5.411	0.919	0.040 ¹	--	1337
Audit	Audit report type	0	1	--	--	--	--	0	2	1337
Profit	Profitability	0	1	--	--	--	--	1	2	1337
CashFlow	Operating cash flow	0	1	--	--	--	--	1	2	1337
TaxCom	Tax compliance	0	1	--	--	--	--	1	2	1337
TaxBurd	Tax burden	0	9	--	--	--	--	9	7	1337
ICTPlan-cof	Comprehensive Tax Plan coefficient	0.013	0.921	0.425	0.342	0.73	0.492	-	--	1337

Table 1 shows that several variables exhibit extreme skewness (e.g., BOBO with a skewness of 18.287, CashHold with 6.829, SMO with 5.411, SOIC with 2.987). Some of the quantitative variables are skewed. Although skewness alone does not necessarily require discretization, this study adopts a tiles-based discretization approach for most quantitative variables for methodological and interpretive reasons. First, the study's primary objective is not parameter estimation but the identification and ranking of tax compliance determinants using decision-tree-based models. Discretization facilitates the extraction of interpretable decision rules and reduces sensitivity to extreme values in highly skewed firm-level variables. Second, tile discretization preserves the relative ordering of observations while creating balanced intervals that are robust to skewness and outliers. To fix this, the Tiles discretization technique was used to turn the Earnings Smoothing (SMO), Disclosure Quality (Disclos), Cash Holding Level (CashHold), Firm Size (Size), State Ownership and Influence in Companies (SOIC), Institutional Ownership Percentage (POIN), Free Float Shares (FRFL), Board Bonus (BOBO), Non-Executive Board Members (OMBS), and the comprehensive tax plan (ICTPlan-cof) variables into ordinal qualitative variables. The fitted models derived from the data are delineated as follows:

4.2. CART Decision Tree

CART is a binary decision tree technique that sequentially partitions the feature space by selecting splits that maximize node purity (homogeneity). The splitting process continues until a stopping rule is satisfied (e.g., minimum node size or insufficient impurity reduction). The algorithm can be applied

in two modes: classification and regression. Given that the dependent variable in this study is categorical (TaxCom), the classification mode of CART was employed. The confusion matrix for the test data is presented in Table 2. To assess whether the observed accuracy is meaningful, we report a simple baseline that always predicts the majority class in the test set. Based on Table 2, the test set contains 293 “compliant” observations (Class 1) and 114 “non-compliant” observations (Class 0), i.e., 72.0% and 28.0%, respectively. Therefore, a majority-class baseline would achieve an accuracy of 0.720 by predicting Class 1 for all cases. In contrast, CART attains an accuracy of 0.54 on the test set. This indicates that overall accuracy alone is not an informative metric in the presence of class imbalance, and the added value of CART should be evaluated using class-specific measures and lift rather than raw accuracy. In particular, while the baseline has zero ability to identify the minority class (Class 0) (Recall₀ = 0.00), CART provides non-trivial identification of the minority class (Recall₀ = 0.52; Table 2), which is critical for detecting non-compliance.

Table 2. Confusion matrix of the CART Algorithm

Actual Values	CART Algorithm Predictions		
	Negative (N)	Positive (P)	
True (T)	59	159	Correct predictions
False (F)	134	55	Incorrect predictions

Equations were used to construct the model evaluation metrics from the confusion matrix results, as shown in Table 2.

$$\begin{aligned}
 Acc &= \frac{(TP + TN)}{(TP + TN + FP + FN)} = \frac{(159 + 59)}{(159 + 59 + 55 + 134)} = 0.54 \\
 Recall^+ &= \frac{TP}{(TP + 0 + 0 + FN)} = \frac{159}{(159 + 0 + 0 + 134)} = 0.54 \\
 Recall^- &= \frac{TN}{(0 + TN + FP + 0)} = \frac{59}{(0 + 59 + 55 + 0)} = 0.52 \\
 Precision^+ &= \frac{TP}{(TP + 0 + FP + 0)} = \frac{159}{(159 + 0 + 55 + 0)} = 0.74 \\
 Precision^- &= \frac{TN}{(0 + TN + 0 + FN)} = \frac{59}{(0 + 59 + 0 + 134)} = 0.31 \\
 F - Measure^- &= \frac{2 * Recall^- * Precision^-}{Recall^- + Precision^-} = \frac{2 * 0.52 * 0.31}{0.52 + 0.31} = 0.388 \\
 F - Measure^+ &= \frac{2 * Recall^+ * Precision^+}{Recall^+ + Precision^+} = \frac{2 * 0.54 * 0.74}{0.54 + 0.74} = 0.624 \\
 Lift &= 1.22
 \end{aligned}$$

As shown in Table 3, the importance of the identified variables and their respective predictive significance levels in the CART decision tree algorithm are presented. Lift measures how much more effectively the decision tree identifies the target (minority) class compared to a random classifier, making it particularly informative in imbalanced classification settings where accuracy alone may be misleading.

Table 3. Predictive importance of variables in CART

No.	Variable	Symbol	predictor importance
1	Non-executive board members	OMBS	0.110
2	Level of cash retention	CashHold	0.110
3	Percentage of institutional shareholder ownership	POIN	0.090
4	Company size	Size	0.080
5	Government ownership and influence in companies	SOIC	0.080
6	Operating cash flow	CashFlow	0.060
7	Disclosure quality	Disclos	0.060
8	Board remuneration	BOBO	0.060
9	Tax burden	taxburd	0.060

The predictor importance values in Table 3 are derived from IBM SPSS Modeler and represent normalized importance scores relative to the most influential predictor, which is scaled to 1. Accordingly, the reported values do not sum to 1 but indicate each variable's relative contribution to impurity reduction in the CART tree relative to the strongest predictor. Table 3 reports the predictor importance values obtained from the CART model. The prominence of non-executive board members is consistent with agency theory, which highlights the monitoring role of independent directors in limiting opportunistic managerial behavior, including aggressive tax practices. The importance of cash holdings reflects liquidity-based explanations of tax compliance, as firms with stronger liquidity positions face fewer constraints in meeting tax obligations. Institutional ownership aligns with corporate governance theory, emphasizing the disciplining role of professional and sophisticated shareholders. The role of firm size captures both higher audit probability and stronger legitimacy pressures faced by larger firms. Finally, the significance of government ownership and influence is consistent with institutional theory, particularly in contexts where state involvement intensifies formal and informal compliance expectations.

The ICTPlan coefficient (ICTPlan-cof), along with variables such as earnings smoothing, audit type, and profitability, was included in the CART estimation. Their absence among the highest-ranked predictors reflects the structure of the CART algorithm, in which the frequency and strength of node splits drive variable importance. These variables primarily affect tax compliance through interaction and contextual mechanisms rather than as primary splitting variables. In particular, ICTPlan-cof represents a macro-level institutional factor whose impact on tax compliance is largely indirect and materializes through firm-level governance, transparency, and incentive structures, explaining its lower standalone importance in the CART ranking.

4.3. C5.0 Decision Tree

The C5.0 decision tree algorithm is similar to the CART algorithm; however, it differs in its pruning strategy. For classification, this approach partitions the dataset into subsets, with records more similar to one another than to the parent node. In the C5.0 decision tree algorithm, samples are split on the field that yields the highest information gain. The first branching determines each subgroup, and then division based on new fields continues until further splitting is impossible. Afterwards, the branches are retested starting at the lowest level, and any that do not make a meaningful contribution to the model are deleted. In Table 4, the confusion matrix for the test data is shown.

Table 4. Confusion matrix of the C5.0 Algorithm

<i>Actual Values</i>	<i>C5.0 Algorithm Predictions</i>	
	<i>Negative (N)</i>	<i>Positive (P)</i>
<i>True (T)</i>	49	202
<i>False (F)</i>	91	65

Table 4 presents the confusion matrix findings used to calculate the following model evaluation criteria. The C5.0 results show asymmetric performance, with higher Precision⁺ and Recall⁺ than Precision⁻ and Recall⁻, indicating that the model identifies compliant firms more accurately than noncompliant firms. This asymmetry reflects the imbalanced nature of tax compliance data and implies that misclassifying noncompliant firms as compliant may be more costly from a policy perspective. Therefore, the model is best interpreted as a screening tool for risk-based audit prioritization rather than a definitive compliance classifier:

$$\begin{aligned}
 Acc &= \frac{(TP + TN)}{(TP + TN + FP + FN)} = \frac{(202 + 49)}{(202 + 49 + 65 + 91)} = 0.62 \\
 Recall^+ &= \frac{TP}{(TP + 0 + 0 + FN)} = \frac{202}{(202 + 0 + 0 + 91)} = 0.69 \\
 Recall^- &= \frac{TN}{(0 + TN + FP + 0)} = \frac{49}{(0 + 49 + 65 + 0)} = 0.43 \\
 Precision^+ &= \frac{TP}{(TP + 0 + FP + 0)} = \frac{202}{(202 + 0 + 65 + 0)} = 0.76 \\
 Precision^- &= \frac{TN}{(0 + TN + 0 + FN)} = \frac{49}{(0 + 49 + 0 + 91)} = 0.35 \\
 F - Measure^- &= \frac{2 * Recall^- * Precision^-}{Recall^- + Precision^-} = \frac{2 * 0.43 * 0.35}{0.43 + 0.35} = 0.385 \\
 F - Measure^+ &= \frac{2 * Recall^+ * Precision^+}{Recall^+ + Precision^+} = \frac{2 * 0.69 * 0.76}{0.69 + 0.76} = 0.723 \\
 Lift &= 1.40
 \end{aligned}$$

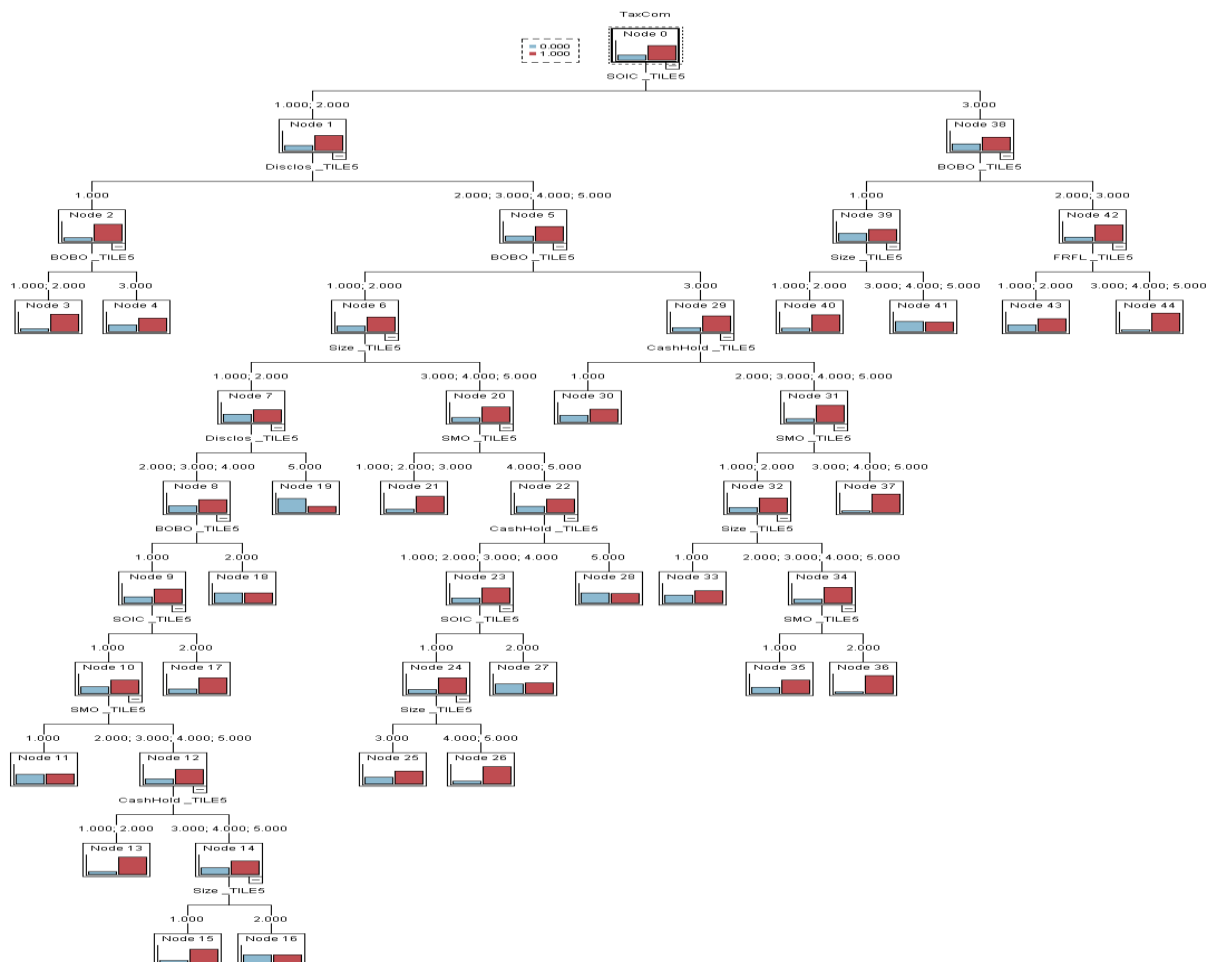
As shown in Table 5, the importance of the variables specified in the table and the degree of importance of predicting the variables in the C5.0 decision tree algorithm are as follows. All conceptually relevant variables were included in the C5.0 model. Variables not reported in Table 5 show negligible normalized importance after tree construction and, therefore, are not retained among the most influential predictors; their absence reflects data-driven selection rather than exclusion from the analysis.

Table 5. Predictive importance of variables in C5.0

No.	Variable	Symbol	predictor importance
1	Disclosure quality	Disclos	0.520
2	Board remuneration	BOBO	0.210
3	Free float percentage	FRFL	0.170
4	Government ownership and influence in companies	SOIC	0.060
5	Cash holding level	CashHold	0.030
6	Profit smoothing	SMO	0.020
7	Company size	Size	0.010

Consequently, in this model, the order of ranking variables is: quality of disclosure, compensation for the board of directors, percentage of free-float ownership, government interest and influence on companies they invest in, cash retention level, profit management smoothing, and company size.

The tree structure of the C5.0 model is also displayed in Figure 2:

**Figure 2.** Tree structure of the C5.0 model

4.4. CHAID Decision Tree

The CHAID (Chi-squared Automatic Interaction Detection) algorithm finds the most important relationships between independent and dependent variables using the p-value from the chi-square statistic in the test of independence for contingency tables. In this approach, the variable with the minimum p-value is chosen first from among those available for splitting at a given node. CHAID differs from other decision tree methods in that it can have more than two branches at any level of the tree. Thus, it is not a binary decision tree like CART and C5.0. The tree structure that CHAID produces is usually wider and more complex than those produced by other methods. Table 6 shows the confusion matrix for the test data used by this program.

Table 6. Confusion matrix of the CHAID Algorithm

Actual Values	CHAID Algorithm Predictions		Positive (P)
	Negative (N)		
True (T)	78	139	Correct predictions
False (F)	154	36	Incorrect predictions

Using the results of the confusion matrix, as shown in Table 6, the model evaluation criteria have been calculated as follows:

$$\begin{aligned}
 Acc &= \frac{(TP + TN)}{(TP + TN + FP + FN)} = \frac{(139 + 78)}{(139 + 78 + 36 + 154)} = 0.53 \\
 Recall^+ &= \frac{TP}{(TP + 0 + 0 + FN)} = \frac{139}{(139 + 0 + 0 + 154)} = 0.48 \\
 Recall^- &= \frac{TN}{(0 + TN + FP + 0)} = \frac{78}{(0 + 78 + 36 + 0)} = 0.68 \\
 Precision^+ &= \frac{TP}{(TP + 0 + FP + 0)} = \frac{139}{(139 + 0 + 36 + 0)} = 0.79 \\
 Precision^- &= \frac{TN}{(0 + TN + 0 + FN)} = \frac{78}{(0 + 78 + 0 + 154)} = 0.36 \\
 F - Measure^- &= \frac{2 * Recall^- * Precision^-}{Recall^- + Precision^-} = \frac{2 * .68 * .36}{.68 + .36} = 0.47 \\
 F - Measure^+ &= \frac{2 * Recall^+ * Precision^+}{Recall^+ + Precision^+} = \frac{2 * 0.48 * 0.79}{.48 + 0.79} = 0.579 \\
 Lift &= 1.34
 \end{aligned}$$

As shown in Table 7, the importance of the identified variables and their corresponding predictive significance levels in the CHAID decision tree algorithm are presented. The dominance of government ownership (SOIC), institutional ownership (POIN), and the comprehensive tax plan (ICTPlan) in the CHAID model underscores the central role of institutional and ownership structures in shaping tax compliance in Iran. Given the strong presence of the state and concentrated ownership, compliance is more closely linked to regulatory oversight and enforcement capacity than to purely firm-level financial characteristics, consistent with institutional approaches to tax compliance.

Table 7. Predictive importance of variables in CHAID

No.	Variable	Symbol	Predictor importance
1	Government ownership and influence in companies	SOIC	0.190
2	Percentage of institutional shareholder ownership	POIN	0.170
3	Iran's comprehensive tax plan	ICTPlan	0.150
4	Profit smoothing	SMO	0.140
5	Disclosure quality	Disclos	0.100
6	Percentage of free float shares	FRFL	0.100
7	Company size	Size	0.080
8	Level of cash retention	CashHold	0.050
9	Non-executive board members	OMBS	0.020
10	Board remuneration	BOBO	0.010

Therefore, in this model, the control variables have been ranked as follows: own state and influence on companies, the institutions' ownership percentage, the Century Iran Management Act, the earnings management smoothing rate, the quality of disclosure, traded shares proportion remaining in free market shares, company size, access to cash, companies level of non-captive board members, as well as executive pay rate.

4.5. QUEST Decision Tree

The QUEST (Quick, Unbiased, Efficient Statistical Tree) algorithm is a new method for expanding binary trees that uses only symbolic output fields and supports only two subgroups at each split in its recursive data partitioning method. The algorithm is efficient in terms of computing and uses backward pruning. Table 8 presents the confusion matrix for the test data collected using the QUEST algorithm.

Table 8. Confusion matrix of the QUEST Algorithm

Actual Values	QUEST Algorithm Predictions		
	Negative (N)	Positive (P)	
True (T)	39	187	Correct predictions
False (F)	106	75	Incorrect predictions

Using the results of the confusion matrix, as shown in Table 8, the model evaluation criteria have been calculated as follows:

$$Acc = \frac{(TP + TN)}{(TP + TN + FP + FN)} = \frac{(187 + 39)}{(187 + 39 + 75 + 106)} = 0.56$$

$$Recall^+ = \frac{TP}{(TP + 0 + 0 + FN)} = \frac{187}{(187 + 0 + 0 + 106)} = 0.64$$

$$\begin{aligned}
 Recall^- &= \frac{TN}{(0 + TN + FP + 0)} = \frac{39}{(0 + 39 + 75 + 0)} = 0.34 \\
 Precision^+ &= \frac{TP}{(TP + 0 + FP + 0)} = \frac{187}{(187 + 0 + 75 + 0)} = 0.71 \\
 Precision^- &= \frac{TN}{(0 + TN + 0 + FN)} = \frac{39}{(0 + 39 + 0 + 106)} = 0.27 \\
 F - Measure^- &= \frac{2 * Recall^- * Precision^-}{Recall^- + Precision^-} = \frac{2 * .34 * .27}{.34 + .27} = 0.30 \\
 F - Measure^+ &= \frac{2 * Recall^+ * Precision^+}{Recall^+ + Precision^+} = \frac{2 * 0.64 * 0.71}{.64 + 0.71} = 0.67 \\
 Lift &= 1.08
 \end{aligned}$$

As shown in Table 9, the variables' importance and predictive importance in the QUEST decision tree algorithm are as follows.

Table 9. Predictive importance of variables in QUEST

No.	Variable	Symbol	predictor importance
1	Non-executive board members	OMBS	0.110
2	Level of cash retention	CashHold	0.110
3	Percentage of institutional shareholder ownership	POIN	0.090
4	Company size	Size	0.080
5	Government ownership and influence in companies	SOIC	0.080
6	Profit smoothing	SMO	0.070
7	Operating cash flow	CashFlow	0.060
8	Disclosure quality	Disclos	0.060
9	Iran's comprehensive tax plan	ICTPlan	0.060
10	Tax burden	taxburd	0.060

As a result, in this algorithm, the variables are prioritized as follows: non-executive board members, cash-holding level, institutional ownership percentage, firm size, state ownership and influence in companies, earnings smoothing, operating cash flow, disclosure quality, Iran's comprehensive tax plan, and tax burden.

4.6. Model comparison

Comparing Models Table 10 presents the findings of the model evaluation comparison in a separate format for each decision tree model. This includes the confusion matrices and assessment metrics.

Table 10. Comparison of model evaluations

Model	TN	FP	FN	TP	Recall-	Recall+	Precision -	Precision +	Lift-	Accuracy
CART	59	55	134	159	0.520	0.540	0.310	0.740	1.220	0.540
CHAID	78	36	154	139	0.680	0.470	0.340	0.790	1.340	0.530
Quest	39	75	106	187	0.340	0.640	0.270	0.710	1.080	0.560
C5.0	49	65	91	202	0.430	0.690	0.350	0.760	1.400	0.620

From the results in Table 10 and the comparison of evaluation measures, especially overall accuracy and Lift Index, the C5.0 model is the best-performing. The C5.0 decision tree resulted in an accuracy rate (ratio of correctly classified observations to the total number of cases in the tax compliance classification) for tax-compliant firms (proportion would refer to the ratio of correctly classified observations over the total number of cases) of 62%, and the quantity predicted by class correctly predicted: 69%. The Precision (nc) of the Recall was 31%, and the Precision was 26%. For the other model class-compliant, non-compliant discretization: random assignment of firms to this class would have a 26% chance of being correct for the conforming class. At the same time, with perfect prediction accuracy (state-of-the-art), 25% was obtained. In contrast, the 36% of non-complying classes as correct by the model suggests that the probability of incorrect failure detection increased by a factor of 1.4 (Lift Index = 1.4).

5. Conclusion and recommendations

The results of the ranking of factors affecting tax compliance, obtained using various decision tree algorithms within the comprehensive tax plan framework, are summarized in Table 11.

According to Table 11, which presents the study's ranking of tax compliance elements, the C5.0 model outperforms the other algorithms examined in model evaluation measures such as overall accuracy and the Lift Index.

Disclosure quality, board remuneration, percentage of free-float shares, cash-holding level, earnings smoothing, firm size, and state ownership and influence in companies are the independent variables that rank most important in shaping the culture of tax compliance, according to the research findings.

This study examined the determinants of corporate tax compliance within the institutional framework of Iran's Comprehensive Tax Plan, using multiple decision-tree-based machine-learning models. The comparison of model performance indicates that the C5.0 algorithm yields the best overall results, achieving the highest accuracy (62%) and lift (1.40), suggesting a meaningful improvement over random classification for identifying noncompliant firms. Beyond predictive performance, the findings offer important theoretical insights. The high importance of disclosure quality and board remuneration is consistent with agency theory, which emphasizes the role of transparency and incentive alignment in mitigating opportunistic behavior, including aggressive tax practices. Similarly, the prominence of free float and institutional ownership highlights the disciplining role of external monitoring mechanisms, aligning with prior corporate governance and tax avoidance literature. The significance of cash holding levels and earnings smoothing reflects firms' liquidity management and reporting discretion as channels through which tax compliance behavior is shaped. Importantly, the relevance of government ownership and influence, as well as the Comprehensive Tax Plan (ICTPlan), underscores the institutional dimension of tax compliance in Iran. These results support frameworks emphasizing tax morale, trust, and enforcement capacity, suggesting that compliance is not driven solely by firm-level characteristics but also by the credibility and reach of tax system reforms.

Table 11. Ranking of factors affecting tax compliance

	Algorithm QUEST		Algorithm CHAID		Algorithm C5.0		Algorithm CART
predictor importance	Variable	predictor importance	Variable	predictor importance	Variable	predictor importance	Variable
0.110	Non-executive board members	0.190	Government ownership and influence in companies	0.520	Disclosure quality	0.110	Non-executive board members
0.110	Level of cash holdings	0.170	Percentage of institutional shareholder ownership	0.210	Board remuneration	0.110	Cash holding level
0.090	Percentage of institutional shareholder ownership	0.150	Iran's comprehensive tax plan	0.170	Free float percentage	0.090	Percentage of institutional shareholder ownership
0.080	Company size	0.140	Profit smoothing	0.060	Government ownership and influence in companies	0.080	Company size
0.080	Government ownership and influence in companies	0.100	Disclosure quality	0.030	Cash holding level	0.080	Government ownership and influence in companies
0.070	Profit Smoothing	0.100	Percentage of free float shares	0.020	Earnings smoothing	0.060	Operating cash flow
0.060	Operating Cash	0.080	Company size	0.010	Company size	0.060	Disclosure quality
0.060	Disclosure Quality	0.050	Level of cash retention			0.060	Board remuneration
0.060	Non-Executive Directors	0.020	Non-executive board members			0.060	Tax burden
0.060	Tax Burden	0.010	Board remuneration				

Overall, the findings demonstrated that tax compliance emerges from the interaction between corporate governance mechanisms and institutional enforcement structures, reinforcing the need for integrated policy approaches under Iran's digital tax transformation. The empirical findings can be interpreted in light of the slippery slope framework and the economic theory of crime. The prominence of disclosure quality and corporate governance variables reflects the role of transparency and accountability, which are associated with trust and perceived fairness. At the same time, the importance of the ICTPlan index and ownership-related variables highlights the role of enforcement

capacity, monitoring, and institutional power, which are central to the probability of detection and compliance decisions in the economic theory of crime.

This study is subject to several limitations that should be considered when interpreting the findings. First, the empirical analysis is based on firms listed on the Tehran Stock Exchange, which, while providing audited and standardized data and ensuring high internal validity, may limit the generalizability of the results. Although the Comprehensive Tax Plan represents a nationwide institutional reform affecting the entire tax system, including infrastructure, digital processes, and enforcement mechanisms, the behavior of small and informal taxpayers may differ due to lower transparency and different compliance incentives. Therefore, the findings should be interpreted in the context of formally listed firms, and caution is warranted when extending them to SMEs and other taxpayer groups.

Second, the study relies on historical data from a specific time period, which may not fully capture recent developments in tax administration and digitalization. Third, the analysis is based on proxy measures for tax compliance and the implementation of the Comprehensive Tax Plan, which, although grounded in theory and prior research, may not fully reflect underlying behavioral dimensions.

Based on these limitations, future research could extend the analysis to SMEs and unlisted entities using administrative or survey-based data, apply alternative machine-learning or econometric approaches, and incorporate direct behavioral measures to further enrich and validate the findings.

References

1. Abdelfattah, T. and Aboud, A. (2020). Tax avoidance, corporate governance, and corporate social responsibility: The case of the Egyptian capital market. *Journal of International Accounting, Auditing and Taxation*, 38, A. 100304. <https://doi.org/10.1016/j.intaccaudtax.2020.100304>
2. Alasfour, F., Samy, M. and Bampton, R. (2016). The determinants of tax morale and tax compliance: Evidence from Jordan. In *Advances in Taxation*, (pp. 125-171). Emerald Group Publishing Limited. <https://doi.org/10.1108/S1058-749720160000023005>
3. Allingham, M. G. and Sandmo, A. (1972). Income tax evasion: A theoretical analysis. *Journal of Public Economics*, 1(3-4), pp. 323-338. [https://doi.org/10.1016/0047-2727\(72\)90010-2](https://doi.org/10.1016/0047-2727(72)90010-2)
4. Alm, J. and Torgler, B. (2011). Do ethics matter? Tax compliance and morality. *Journal of Business Ethics*, 101(4), pp. 635-651. <https://doi.org/10.1007/s10551-011-0761-9>
5. Armstrong, C. S., Blouin, J. L., Jagolinzer, A. D. and Larcker, D. F. (2015). Corporate governance, incentives, and tax avoidance. *Journal of Accounting and Economics*, 60(1), pp. 1-17. <https://doi.org/10.1016/j.jacceco.2015.02.003>
6. Balakrishnan, K., Blouin, J. L. and Guay, W. R. (2019). Tax aggressiveness and corporate transparency. *The Accounting Review*, 94(1), pp. 45-69. <https://doi.org/10.2308/accr-52130>
7. Bani-Khalid, T., Alshira'h, A. F. and Alshirah, M. H. (2022). Determinants of tax compliance intention among Jordanian SMEs: A focus on the theory of planned behavior. *Economies*, 10(2), p. 30. <https://doi.org/10.3390/economies10020030>
8. Bartha, A. and Boda, Z. (2024). Tax compliance motivations during corruption scandals in a fragile democracy: a before-and-after study. *Europe-Asia Studies*, 76(7), pp. 1055-1077. <https://doi.org/10.1080/09668136.2024.2378814>
9. Block, J. (2016). Körperschaftsteuer und Unternehmensgründung. IZA World of Labor, Bonn, Germany, p. 257. <https://doi.org/10.15185/izawol.257>
10. Brockmeyer, A., Estefan, A., Arras, K. R. and Serrato, J. C. S. (2021). Taxing property in

- developing countries: Theory and evidence from Mexico (No. w28637). National Bureau of Economic Research. Working Papers, United States. <https://doi.org/10.3386/w28637>
11. De Neve, J. E., Imbert, C., Spinnewijn, J., Tsankova, T. and Luts, M. (2021). How to improve tax compliance? Evidence from population-wide experiments in Belgium. *Journal of Political Economy*, 129(5), pp. 1425-1463. <https://doi.org/10.1086/713096>
 12. Desai, M. A. and Dharmapala, D. (2006). Corporate tax avoidance and high-powered incentives. *Journal of Financial Economics*, 79(1), pp. 145-179. <https://doi.org/10.1016/j.jfineco.2005.02.002>
 13. Dikmen, S., Çiçek, H. G. and Kaya, I. (2025). How does Spiritual Well-Being and Religious attitude influence tax compliance? Evidence from Türkiye. *Journal of Tax Reform*, 11(2), pp. 417-433. <https://doi.org/10.15826/jtr.2025.11.2.209>
 14. Duhoon, A. and Singh, M. (2023). Corporate tax avoidance: a systematic literature review and future research directions. *LBS Journal of Management & Research*, 21(2), pp. 197-217. <https://doi.org/10.1108/LBSJMR-12-2022-0082>
 15. Dyreng, S. D., Hanlon, M. and Maydew, E. L. (2008). Long-run corporate tax avoidance. *The Accounting Review*, 83(1), pp. 61-82. <https://doi.org/10.2308/accr.2008.83.1.61>
 16. Foley, C. F., Hartzell, J. C., Titman, S. and Twite, G. (2007). Why do firms hold so much cash? A tax-based explanation. *Journal of Financial Economics*, 86(3), pp. 579-607. <https://doi.org/10.1016/j.jfineco.2006.11.006>
 17. Gebrihet, H. G., Gebresilassie, Y. H. and Woldu, G. T. (2023). Trust, corruption, and tax compliance in fragile states: on a quest for transforming Africa into future global powerhouse. *Social Sciences*, 13(1), p. 3. <https://doi.org/10.3390/socsci13010003>
 18. Hesami, S., Jenkins, H. and Jenkins, G. P. (2024). Digital transformation of tax administration and compliance: A systematic literature review on E-Invoicing and prefilled returns. *Digital Government: Research and Practice*, 5(3), pp. 1-20. <https://doi.org/10.1145/3643687>
 19. Hsu, H. Y. (2024). Fiscal transparency and tax morale: Is the relationship shaped by perceptions of government performance and corruption?. *International Review of Administrative Sciences*, 90(3), pp. 563-580. <https://doi.org/10.1177/00208523231220599>
 20. International Monetary Fund (IMF). (2022). Evidence from VAT E-Invoicing Reform in Peru, IMF Working Paper, Washington, DC. <https://doi.org/10.5089/9798400204876.001>
 21. Kaghazloo, F. and Borrego, A. C. (2022). Designing a model of the factors affecting tax professionals' tax noncompliant behaviour using the ISM Approach. *Public Organization Review*, 22(4), pp. 1099-1120. <https://doi.org/10.1007/s11115-021-00572-y>
 22. Kerr, J. N., Price, R. and Román, F. J. (2016). The effect of corporate governance on tax avoidance. In Proceedings. *Annual Conference on Taxation and Minutes of the Annual Meeting of the National Tax Association*, National Tax Association, 109, pp. 1-50. <https://dx.doi.org/10.2139/ssrn.3861965>
 23. Kirchler, E., Hoelzl, E. and Wahl, I. (2008). Enforced versus voluntary tax compliance: The “slippery slope” framework. *Journal of Economic Psychology*, 29(2), pp. 210-225. <https://doi.org/10.1016/j.joep.2007.05.004>
 24. Koumpias, A. M., Leonardo, G. and Martinez-Vazquez, J. (2020). Trust in government institutions and tax morale. International Center for Public Policy, Andrew Young School of Policy Studies, Georgia State University. Working Paper, Atlanta, Georgia <https://icepp.gsu.edu/files/2020/02/paper2001.pdf>
 25. Mebratu, A. A. (2024). Theoretical foundations of voluntary tax compliance: evidence from a developing country. *Humanities and Social Sciences Communications*, 11(1), pp. 1-8. <https://doi.org/10.1057/s41599-024-02903-y>

26. Mohammed, H. and Tangl, A. (2023). Taxation perspectives: Analyzing the factors behind viewing taxes as punishment—A comprehensive study of taxes as service or strain. *Journal of Risk and Financial Management*, 17(1), p. 5. <https://doi.org/10.3390/jrfm17010005>
27. Nwokoye, E. S., Igbunugo, C. I., Ekesiobi, C. and Dimnwobi, S. K. (2023). Fiscal incentives and tax compliance behaviour in industrial clusters: a survey of clusters in South-east Nigeria. *Journal of African Business*, 24(1), pp. 147-166. <https://doi.org/10.1080/15228916.2022.2031827>
28. Opler, T., Pinkowitz, L., Stulz, R. and Williamson, R. (2001). Corporate cash holdings. *Journal of Applied Corporate Finance*, 14(1), pp. 55-67. <https://doi.org/10.1111/j.1745-6622.2001.tb00320.x>
29. Organisation for Economic Co-operation and Development (OECD). (2024). Tax Administration 2024: Insights from Digital Transformation, OECD Publishing, Paris. <https://doi.org/10.1787/2d5fba9c-en>
30. Rwabunywenge, M., Nchimbi, M. and Suluo, S. (2024). The influence of provision of public services on tax compliance in SMEs: mediating role of tax morale. *Business Management Review*, 27(2), pp. 90-107. <https://doi.org/10.56279/bmr.v27i2.6919>
31. Sano, H. (2025). Impacts of public disclosure on tax compliance using agent-based modeling: H. Sano. *Journal of Economic Interaction and Coordination*, 20(1), pp. 273-305. <https://doi.org/10.1007/s11403-024-00420-4>
32. Slemrod, J. (2019). Tax compliance and enforcement. *Journal of Economic Literature*, 57(4), pp. 904-954. <https://doi.org/10.1257/jel.20181437>
33. Sokolova, M. and Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427-437. <https://doi.org/10.1016/j.ipm.2009.03.002>
34. Taing, H. B. and Chang, Y. (2021). Determinants of tax compliance intention: Focus on the theory of planned behavior. *International Journal of Public Administration*, 44(1), pp. 62-73. <https://doi.org/10.1080/01900692.2020.1728313>
35. Torgler, B. (2007). Tax compliance and tax morale: A theoretical and empirical analysis. In *Tax Compliance and Tax Morale*. Edward Elgar Publishing, Massachusetts, USA.
36. Zhang, L., Gu, Y., Yu, Y. and Pan, L. (2025). Mandatory industry disclosure and corporate tax avoidance: Evidence from Chinese listed firms. *Economic Analysis and Policy*, 109, A. 101469. <https://doi.org/10.1016/j.eap.2025.12.008>



RESEARCH ARTICLE

Greener Growth, Tighter Credit? The Dual Role of Climate Action in Revenue Expectations and Financing Obstacles

Ly Hoang Vu

University of Finance - Marketing, Ho Chi Minh City, Vietnam

How to cite this article:

Vu, L. H. (2026). Greener Growth, Tighter Credit? The Dual Role of Climate Action in Revenue Expectations and Financing Obstacles. *Iranian Journal of Accounting, Auditing and Finance*, 10(3), 47-77. doi: 10.22067/ijaaf.2026.48267.1662
https://ijaaf.um.ac.ir/article_48267.html

ARTICLE INFO

Article History

Received: 2025-12-06

Accepted: 2026-02-03

Published online: 2026-07-12

Keywords:

Climate Change Mitigation,
Access to Finance, Female
Leadership, Female
Ownership, Revenue
Expectations

Abstract

This study examines the firm-level determinants of expected revenue growth and perceived ease of access to finance, with a particular focus on climate change mitigation actions, leadership and ownership characteristics, institutional environments and macroeconomic conditions. The analysis of 21,729 enterprises across 28 countries during 2018–2020 showed that firms that adopt climate change mitigation actions are significantly more likely to expect revenue growth. Second, female leadership was positively associated with revenue growth expectations, whereas a higher share of female ownership was linked to lower growth expectations. Third, higher perceived corruption and inflation were associated with reduced growth expectations. Regarding perceived access to finance, larger firms, those with more experienced top managers, and firms operating in less corrupt environments perceived easier access to finance. Conversely, firms that monitor CO₂ emissions and those that have experienced weather-related losses report significantly greater financing obstacles. These findings suggested that climate mitigation efforts are not environmentally beneficial alone but are positively correlated with managerial optimism about future revenue. However, such efforts, particularly emissions monitoring, may also signal risks to financiers, while climate-related physical shocks further constrain perceived access to capital. The results underscored the importance of both firm-level capabilities and institutional quality in shaping expectations and financial constraints.


<https://doi.org/10.22067/ijaaf.2026.48267.1662>


NUMBER OF REFERENCES
45



NUMBER OF FIGURES
0



NUMBER OF TABLES
14

Homepage: <https://ijaaf.um.ac.ir>
 E-ISSN: 2717-4131
 P-ISSN: 2588-6142

*Corresponding Author: Ly Hoang Vu
 Email: lyhoangvu@ufm.edu.vn
 Tel:
 ORCID: 0000-0003-4948-8863

1. Introduction

Economic growth, employment creation, and technical progress are all fueled by expanding businesses. Many things influence a company's course, but two of the most essential are the likelihood of future revenue growth and the perceived ease of access to capital. Investment choices, recruitment, and innovation are influenced by managers' optimism and strategic, forward-looking behavior, as reflected in revenue growth expectations. Companies' ability to secure working capital, expand their operations, and withstand economic shocks results from their growth ambitions, which, in turn, are influenced by their perceptions of access to finance. The capacity to anticipate growth and the certainty that external funding will be available when required are essential for businesses of any size to grow sustainably. Despite a wealth of literature on firm-level performance, even less is known about the interplay between these two crucial dimensions and emerging strategic priorities, particularly the mitigation of climate change, as well as the moderating effects of leadership and ownership traits, institutional quality, and macroeconomic circumstances.

There has been a lot of momentum in recent years for the economic rationale of climate change action. Regulatory demands, stakeholder expectations, or genuine environmental commitment are driving businesses to adopt techniques such as energy management systems and CO₂ emission monitoring. Nevertheless, there is ongoing debate about the relationship between these climate mitigation efforts and business success. One side of the argument states that energy efficiency should increase revenue expectations and make credit easier to get because it lowers operational costs, improves reputational capital, and signals forward-looking management. However, climate initiatives require upfront investment, and monitoring emissions could put firms at risk of future regulations or costly transitions, which could make lenders wary or reduce optimism. Creditworthiness and the ability to attract capital can be improved through proactive mitigation measures such as monitoring CO₂ emissions or improving energy efficiency (Eccles and Klimenko, 2019). On the other hand, if we do nothing to address environmental issues, we risk seeing capital costs rise or our financing options decline (Flammer, 2021). The importance of financial structure in promoting low-carbon growth is supported by research from across countries, as presented by De Haas and Popov (2023). They conclude that broader stock markets accelerate emissions reductions in carbon-intensive industries by enabling green innovation and the reallocation of investment towards more energy-efficient sectors. The panel covers 48 countries, 16 industries, and 26 years. The importance of financial development, and the rise of equity markets in particular, in facilitating the shift to sustainable growth is highlighted by their findings. To reduce emissions and increase resource efficiency, businesses must take the lead in adopting mitigation strategies, even while governments establish legal frameworks (Porter and Van Der Linde, 1995). Stefan and Paul (2008) and Peters and Romi (2015) are two studies that examined the effects of firm size, sector, governance, and leadership on environmental behavior. But there are still three major holes. Despite evidence that managerial and ownership roles vary in power and strategic aims, the unique impact of gender on each is difficult to separate. The second issue is that in multi-country research, the impact of macroeconomic factors, such as inflation and average GDP levels, on the adoption of mitigation strategies has largely been ignored. Thirdly, there has been a lack of systematic research on the financial access conundrum as it pertains to climate action, wherein actual and perceived credit access differ.

For these conflicting ideas, there is a dearth of empirical evidence, especially from large-scale, cross-national business surveys. This study fills that need by investigating the potential correlation between two climate actions CO₂ monitoring and energy management and improved perceived access to financing and more accurate revenue growth predictions. Paying close attention to leadership and ownership traits, particularly gender composition, is essential for understanding firm-level expectations and climate initiatives. Women business owners and managers may have different risk tolerances, strategic orientations, and network access than men, according to a growing corpus of

research. While some research shows that companies run by women are more environmentally conscious or more profitable, other studies reveal that they are less willing to take financial risks or face discriminatory lending practices. There is a lack of large-sample cross-national evidence on how female ownership and leadership influence growth expectations and perceptions of finance, and the empirical evidence is inconsistent. This study makes a valuable contribution by methodically examining the potential impact of female top management or a higher percentage of female ownership on these outcomes. It does this while controlling for a comprehensive set of business and environmental characteristics.

Institutional context also matters profoundly. Corruption, defined as the abuse of public power for private gain, is widely recognized as a binding constraint on firm activity. In corrupt environments, firms may face bribe demands for loan approvals, opaque regulatory enforcement, or unreliable contracts. Such conditions are likely to depress managerial optimism about future growth and severely impair perceived access to finance. Yet, the magnitude of corruption's effect relative to other firm-level and macroeconomic factors has not been extensively quantified across diverse national contexts. This study provides robust estimates of how perceived corruption, as an operational obstacle, affects both expected revenue growth and access to financing. Macroeconomic conditions, particularly GDP and inflation, constitute the broader environment within which firms operate. Conventional wisdom holds that stronger economic growth boosts firm-level revenue expectations, while high inflation erodes purchasing power and creates uncertainty, dampening optimism and tightening credit availability. However, existing research has primarily examined these relationships using aggregate data or small samples. Whether macroeconomic averages over five- or ten-year horizons exert independent effects on firm-level perceptions, after controlling for firm characteristics and institutional factors, remains an open empirical question. This study tests whether long-term GDP and inflation are significantly associated with expected revenue growth and perceived access to financing, or whether micro-level and institutional factors dominate these associations. Finally, firm-level structural characteristics such as size, age, legal status, and sector are well-established correlates of performance and financial access. Larger firms typically enjoy economies of scale, greater collateral, and more formal relationships with financial institutions. Older firms may benefit from reputation and accumulated experience but might also face lower growth ceilings or organizational inertia. Publicly listed firms and manufacturing enterprises may operate under distinct regulatory and competitive environments. This study controls for these factors to isolate the unique contributions of climate actions, gender, corruption, and macroeconomic conditions.

To investigate those relationships, two research questions were posed: “What is the correlation between the firm’s climate actions and its expected revenue growth?” and “What is the correlation between the firm’s climate actions and its perceived ease of accessing finance?” To do so, this study analyzes a unique, large-scale dataset: Round VI of the Business Environment and Enterprise Performance Survey (BEEPS VI), conducted by the European Bank for Reconstruction and Development (EBRD), merged with World Bank Open Data on macroeconomic indicators. BEEPS VI covers approximately 28,000 firms across 41 countries between 2018 and 2020, with detailed information on climate practices, manager and owner characteristics, perceptions of corruption, and operational environments. After merging with macroeconomic data and applying sample restrictions, the analytical sample comprises 21,729 enterprises across 28 countries.

The analysis proceeds in two main parts. First, we examine the correlation between the firm’s climate actions and its expected revenue growth for the following year. Second, we do the same to analyze the relationship between the firm’s climate actions and its perceived ease of accessing finance. By comparing results across these two outcomes, the study illuminates whether factors that drive growth expectations operate similarly to or differently from those that shape perceptions of

financing, a distinction with important implications for both managerial practice and public policy.

The remainder of the paper is structured as follows. Section 2 discusses related literature and hypothesis development. Section 3 describes the empirical strategy, data sources, and variable construction. Section 4 presents the empirical results and robustness checks and addresses endogeneity problems. Section 5 discusses the findings and limitations. Section 6 concludes with important implications.

2. Literature review

2.1. Firm climate change mitigation and expected revenue growth

Energy efficiency measures, the use of renewable energy sources, and emissions monitoring are all part of climate change mitigation techniques that firms can implement (Porter and Van Der Linde, 1995; Stefan and Paul, 2008). Compliance with regulations, cost savings, stakeholder pressure, and the desire to improve one's reputation are common motivators for implementing such measures (Berrone et al., 2012). Because of variations in pricing structures, institutional pressures, and resource availability, patterns of adoption vary greatly among industries and continents (Kolk and Pinkse, 2005; Hoffman, 2007).

Whether climate mitigation activities are linked to firms' projections of future revenue growth is a crucial question for this study. Theoretical viewpoints offer contradictory predictions. Energy management and emissions monitoring, according to the resource-based view, can increase managerial optimism for future sales by reducing operating expenses, improving efficiency, and building reputational capital (Porter and Van Der Linde, 1995; Stefan and Paul, 2008). However, such measures require an initial capital outlay and may prompt future regulatory scrutiny, which could reduce expectations for immediate expansion. There is a lack of large-scale international studies that provide empirical evidence for this connection. After adjusting for business characteristics, gender, corruption, and macroeconomic conditions, this study directly tests whether enterprises that monitor CO₂ emissions or employ energy management strategies report better-predicted revenue growth.

Manufacturing companies are more prone to take action to reduce their environmental impact, according to research (Littlewood et al., 2018). On the other hand, larger, more established companies have more money to invest in green technology (Geroski, 2000). However, SMEs face specific challenges, particularly regarding capital, that can hinder their participation in sustainability efforts (Boso et al., 2013). The empirical analysis uses these patterns as control variables.

2.2. Gender, leadership, and firm performance expectations

There is mounting evidence that shows how environmental and social responsibility are correlated with gender diversity in leadership roles. Greater CSR engagement, stronger environmental disclosure, and more sustainable decision-making are generally related to female involvement in ownership and governance. According to Adams and Funk (2012), female leaders are more likely to prioritize social and environmental goals over financial ones. By extension, this line of thinking suggests that female leadership may impact expectations for future performance, even if there is little evidence to support this claim.

Gender may have a different impact on executive and ownership positions, though. According to Glass et al. (2016), management responsibilities can be constrained by board directives or organizational culture, whereas ownership positions typically entail the authority to allocate capital and control strategic decisions. Female directors can influence the adoption of environmental policies, according to Post and Byron (2015), but only when the organizational structure supports their authority. This difference has not been thoroughly investigated in cross-national settings, particularly in developing economies with diverse gender norms and institutional obstacles. To determine whether

female top management and female ownership percentages have distinct effects on revenue growth expectations and perceived access to financing, this study disaggregates them.

2.3. Macroeconomic environment: Real GDP growth and inflation

The incentives, capacities, and performance expectations of enterprises regarding environmental investment are influenced by the macroeconomic context. According to [Dasgupta et al. \(2002\)](#) and [Bowen et al. \(2012\)](#), there is a correlation between higher GDP levels and improvements in environmental consciousness, technological capacity, and institutional support for sustainability. Prosperity enables investments in resilience and low-carbon transitions over the long run, as [Hallegatte and Rozenberg \(2017\)](#) point out. On the other side, environmental concerns may take a back seat to economic uncertainty ([Barrett and Graddy, 2000](#)). Managerial optimism about firm-level revenue estimates should be higher in an economy with faster GDP growth and lower in one with weaker growth. However, there is a lack of data on whether average long-run GDP growth affects firm-level opinions, even after accounting for firm traits, corruption, and climate initiatives.

An important consideration is inflation, which, according to [Fisher \(1930\)](#) and [Bernanke \(1983\)](#), can discourage capital-intensive projects such as energy-efficiency upgrades and emissions monitoring by increasing uncertainty and shortening investment horizons. Managers may start prioritizing short-term cash flow over long-term sustainability targets if inflationary pressures persist ([Aghion et al., 2010](#)). Budgetary instability is a common cause of inflation. If the government's finances are precarious, it will not provide the necessary support for climate mitigation. A decline in purchasing power and an increase in uncertainty and pessimism could be detrimental to anticipated revenue growth if inflation is too high. If lenders become more risk-averse due to inflation, the supply of credit could be limited, affecting perceptions of access to finance. Looking at GDP and inflation averages over five and ten years, this analysis determines whether there is a meaningful relationship between the two variables.

2.4. Access to finance

Access to finance is a critical enabler of sustainability transitions and firm growth. Firms engaging in climate-positive behavior can benefit from lower financing costs, improved investor confidence, and eligibility for green finance instruments ([Goyal and Kumar, 2021](#)). Financial institutions are increasingly integrating ESG metrics into lending decisions ([Kölbel et al., 2017](#)), and empirical evidence shows that sustainability performance correlates with higher credit ratings ([Chava, 2014](#)). If climate actions signal responsible management, firms adopting such measures should perceive easier access to finance.

2.5. Hypothesis development

2.5.1. Climate change mitigation actions and expected revenue growth

The relationship between corporate environmental practices and firm performance has received substantial attention in the literature. The resource-based view (RBV) of the firm suggests that environmental capabilities can become sources of competitive advantage when they are valuable, rare, and inimitable ([Barney, 1991](#); [Hart, 1995](#)). Firms that adopt climate change mitigation actions such as monitoring CO₂ emissions and implementing energy management measures may reduce operational costs, enhance resource efficiency, and preemptively comply with evolving regulations. These improvements directly boost profitability and, by extension, managerial confidence in future revenue growth.

Moreover, signaling theory suggests that voluntary environmental actions convey positive information to stakeholders about management quality and long-term strategic orientation ([Connelly](#)

et al., 2011). A firm that monitors its emissions signals that it is forward-looking, transparent, and prepared for transition risks. Similarly, energy management signals operational excellence and cost consciousness. Both signals should enhance managers' expectations for future sales, as they reflect internal capabilities likely to pay off in the marketplace. We therefore hypothesize:

H1: Firms that adopt climate change mitigation actions are positively associated with expected revenue growth.

2.5.2. Climate change mitigation actions and perceived access to finance

The relationship between climate actions and access to finance is more complex and potentially more ambiguous. On the one hand, the sustainable finance literature suggests that strong environmental practices reduce information asymmetry and signal creditworthiness. Lenders are increasingly incorporating ESG criteria into credit risk assessments, and firms with robust environmental management may receive preferential loan terms (Dhaliwal et al., 2011; Eccles and Klimenko, 2019). Energy management, in particular, directly lowers operating costs, freeing up cash flow for debt service and improving credit metrics. From this perspective, both CO₂ monitoring and energy management should be linked to improved access to finance.

On the other hand, when firms begin measuring their carbon footprint, they may uncover substantial climate-related risks: dependence on fossil-fuel inputs, energy-intensive production processes, exposure to carbon taxes, or supply chain disruptions from extreme weather events. In the absence of mandatory disclosure, many firms avoid monitoring and thus never have to report or confront their emissions. However, once monitoring begins, firms become aware of their vulnerabilities and may communicate these risks, directly or indirectly, to lenders. Risk-averse commercial banks, especially those in emerging markets with limited experience in climate risk assessment, may respond by tightening credit terms, demanding higher collateral, or rejecting loan applications.

Thus, a negative association can arise not because monitoring causes worse financial outcomes, but because it reveals pre-existing risks that lenders then price into credit decisions. This mechanism is grounded in the literature on asymmetric information and environmental risk (Barth et al., 2008; Chava, 2014). Banks are known to penalize firms with poor environmental profiles, but the act of measurement itself can act as a disclosure event. For example, research on the EU Emissions Trading Scheme shows that firms that first reported their emissions experienced stock price declines not because emissions changed, but because previously hidden information was revealed (Griffin et al., 2017). We hypothesize:

H2: Adopting climate change mitigation actions is negatively associated with perceived ease of access to finance.

2.5.3. The moderating role of corruption on female leadership: revenue growth vs. access to finance

The performance implications of female leadership depend critically on institutional context. Drawing on institutional voids theory (Khanna and Palepu, 2010), we argue that corruption, a severe institutional void, produces opposing effects on two key outcomes: expected revenue growth and perceived access to finance.

In corrupt environments, female leaders are often excluded from male-dominated informal networks (Ibarra, 1992; McGuire, 2002). This exclusion, rather than being purely disadvantageous, forces them to develop resilient internal substitutes: transparent accounting, formal compliance, and merit-based practices. Combined with higher risk aversion and lower tolerance for unethical behavior (Adams and Funk, 2012; Croson and Gneezy, 2009), these capabilities improve operational reliability

and reduce exposure to bribery-related volatility. Selection effects may also operate, whereby only the most capable female managers rise to top positions in highly corrupt settings (Ryan and Haslam, 2005). Consequently, female leadership is associated with stronger revenue growth expectations in high-corruption environments.

However, the effect on perceived access to finance operates in the opposite direction. Bank loan officers in corrupt environments are embedded in the same male-dominated networks. Loan approvals are often discretionary and require informal facilitation or bribes (Rajan and Zingales, 1998; Allen et al., 2005). Female executives, lacking these connections, are systematically disadvantaged. Moreover, the transparency and formalization that drive operational resilience may be penalized by corrupt lenders, who view such firms as "difficult" or "rigid." In clean, rule-based environments, this disadvantage disappears, and female leaders may even benefit from their firms' better governance. Thus, a paradoxical pattern emerges: female leaders perform better on growth expectations but worse on access to financing in corrupt settings. We hence pose the following hypotheses:

H3: The association between female top managers and expected revenue growth is moderated by perceived corruption, such that the positive association is stronger in high-corruption environments.

H4: The association between female top managers and perceived ease of accessing finance is moderated by perceived corruption, such that the positive association in low-corruption environments weakens or reverses in high-corruption environments.

3. Methodology and data description

3.1. Methodology

To answer the two research questions, Models 1 and 2 were adopted for expected growth and perceived access to finance, respectively. Ordinary Least Squares (OLS) was applied as the method. Besides that, we also weighted by the total population of the country where the firm is based.

$$REVINCRI_i = \alpha + \beta_1 ACTION_i + \beta_2 F_OWN_i + \beta_3 F_TMAGER_i + \beta_4 TMANAGER_EXP_i + \beta_5 FIRM_i + \beta_6 COUNTRY_i + \mu_i + FE (Area) + FE (Year) \text{ (Model 1)}$$

$$FACCESS_i = \alpha + \beta_1 ACTION_i + \beta_2 F_OWN_i + \beta_3 F_TMAGER_i + \beta_4 TMANAGER_EXP_i + \beta_5 FIRM_i + \beta_6 COUNTRY_i + \mu_i + FE (Area) + FE (Year) \text{ (Model 2)}$$

where

- $REVINCRI_i$ refers to whether the firm i expects its total sales to increase in the following year.
- $FACCESS_i$ relates to the firm i 's perceived ease of accessing finance.
- $ACTION_i$ is climate change mitigation actions.
- F_OWN_i is the percentage of female ownership for firm i
- F_TMAGER_i refers to whether firm i has a female top manager.
- $TMANAGER_EXP_i$ is a variable for the experience of the top manager.
- $FIRM_i$ is a set of control variables for the firm i 's characteristics.
- $COUNTRY_i$ is a set of control variables for the country where firm i is located.
- μ_i is an error term.
- $FE (Area)$ are the geographical area's fixed effects.
- $FE (Year)$ are year fixed effects.

3.2. Data description

The dataset draws on two complementary sources: Round VI of the Business Environment and Enterprise Performance Survey (BEEPS VI), conducted by the European Bank for Reconstruction

and Development (EBRD), and World Bank Open Data. BEEPS VI gathered responses from approximately 28,000 firms across 41 countries between 2018 and 2020, using face-to-face interviews with managers. The survey captures detailed information on firm-level access to credit, expected revenue growth, climate change mitigation actions, owner and manager characteristics, and the broader operational environment. Merging these sources yields a rich and diverse sample of 21,729 enterprises across 28 countries¹, covering the period 2018–2020. The following outlines how these data are used to construct the variables in our analysis.

3.2.1. Expected Revenue Growth (REVINCR)

The variable (REVINCR) captures whether a firm expects its total sales to increase in the following year. It is constructed from the BEEPS VI question: “Considering the next year, are this establishment’s total sales expected to increase, decrease, or stay the same?” Responses are coded as 1 for “Increase” and 0 for “Decrease” or “Stay the same.”

3.2.2. Perceived Ease of Access to Finance (FACCESS)

The variable (FACCESS) measures the firm’s perception of how easily it can access finance. It uses the BEEPS VI question: “To what degree is access to finance an obstacle to the current operations of this establishment?” Possible answers range from “No obstacle” to “Very severe obstacle.” We code responses as follows: 0 for “Very severe obstacle”, 1 for “Major obstacle”, 2 for “Moderate obstacle”, 3 for “Minor obstacle”, 4 for “No obstacle.” Higher values indicate that the firm perceives access to finance as less of an obstacle, implying easier access to financing.

3.2.3. Climate Change Mitigation Actions (ACTION)

Two dummy variables are created based on BEEPS VI questions about firms’ climate change mitigation actions:

- *CO2MONITOR*: Derived from “Over the last three years, did this establishment monitor its CO₂ emissions?” It is coded as 1 for “Yes,” 0 for “No.”
- *ENERGYM*: Derived from the question “Over the last three years, did this establishment adopt energy management measures?” It is coded as 1 for “Yes,” 0 for “No.”
- Together, these variables indicate whether a firm has engaged in energy management or emissions monitoring, both central to climate change mitigation.

3.2.4. Female Leadership and Ownership (F_OWN, F_TMAGER and TMANAGER_EXP)

For female leadership and ownership, we create the following based on answers to questions of the BEEPS VI:

- *F_OWN*: The percentage of the firm owned by females, based on the question: “What percentage of the firm is owned by females?”
- *F_TMAGER*: A dummy variable coded as 1 if the top manager is female and 0 if male. Gender has been shown to influence managerial decision-making, particularly regarding sustainability and innovation (Adams and Funk, 2012). The variable is constructed from the question “Is the Top Manager female?”
- *TMANAGER_EXP*: The number of years of experience the top manager has in the industry, logged to account for diminishing returns. More experienced managers may bring greater

¹ 28 countries include Albania; Azerbaijan; Belarus; Bosnia and Herz.; Bulgaria; Croatia; Czechia; Egypt; Georgia; Greece; Hungary; Italy; Jordan; Kazakhstan; Kyrgyzstan; Lebanon; Lithuania; Poland; Portugal; Romania; Russia; Serbia; Slovak Rep.; Slovenia; Tajikistan; Tunisia; Turkiye; and Ukraine.

expertise and networks, positively influencing climate mitigation and financial access (Hambrick and Mason, 1984; Boone and Hendriks, 2009). This variable is derived from the question “How many years of experience working in this sector does the Top Manager have?” and then log-transformed.

3.2.5. Firm-Level Control Variables (FIRM)

Several firm-specific characteristics are included to account for factors that may influence both expected revenue growth and perceived access to finance:

- *LNREVENUE*: The natural logarithm of last year’s total sales, based on the question: “In the last fiscal year, what were this establishment’s total annual sales for all products and services?”
- *LNAGE*: The natural logarithm of the firm’s age, calculated as the survey year minus the year the firm began operations (from the question “In what year did this establishment begin operations?”).
- *LISTED*: A dummy variable indicating whether the firm is publicly listed, based on the question “What is this firm’s current legal status?” It is coded as 1 for “Shareholding Company with shares traded in the stock market” and 0 for all other legal statuses (e.g., sole proprietorship, partnership, non-traded shares).
- *MSME*: A dummy variable for micro, small, or medium-sized enterprises. It is coded as 1 for “Micro,” “Small,” or “Medium” and 0 for “Large.” SMEs often face greater resource constraints and barriers to financing (Boso et al., 2013).
- *MANUF*: A dummy variable for manufacturing firms, coded as 1 for “Manufacturing” and 0 for other sectors (e.g., services, wholesale/retail, construction). Manufacturing firms typically have larger environmental footprints and distinct financial needs.
- *WEATHER*: A dummy variable indicating whether the firm has experienced monetary losses due to extreme weather events (e.g., storms, floods, droughts, landslides) over the last three years. Based on the corresponding survey question, coded as 1 for “Yes,” 0 for “No.”

3.2.6. Country-Level Control Variables (COUNTRY)

The following country-level variables are included to account for broader economic conditions:

- *GG1Y, GG5Y, GG10Y*: One-year, five-year, and ten-year average real GDP growth rates of the country where the firm operates.
- *INF1Y, INF5Y, INF10Y*: One-year, five-year, and ten-year average inflation rates of the country where the firm is based.

Why do we choose five-year and ten-year averages? Top managers update their beliefs about future sales and financing conditions using all available information, including their accumulated experience of past macroeconomic volatility and long-run trends. Drawing on experience-based learning (Malmendier and Nagel, 2011; Malmendier and Shen, 2024), top managers who have lived through decades of stable growth or high inflation form different expectations from those who have experienced repeated recessions, even when current conditions are identical. Moreover, macroeconomic regimes exhibit substantial persistence, meaning that 5-year and 10-year averages capture deep structural features, such as institutional quality and the investment climate, that shape firms’ long-term planning horizons. For these reasons, long-term averages are theoretically appropriate and often more informative than volatile short-term measures for explaining firm-level expectations and perceptions. In addition to long-term averages, we also include one-year averages for both GDP growth and inflation as an alternative specification.

Those are obtained from the World Bank Open Data and the IMF Data.

CORRUPT: Perception of corruption as an obstacle, based on the BEEPS VI question: “To what degree is corruption an obstacle to the current operations of this establishment?” Responses are coded as: 4 for “Very severe obstacle”, 3 for “Major obstacle”, 2 for “Moderate obstacle”, 1 for “Minor obstacle”, 0 for “No obstacle.” Higher values indicate a more severe perceived obstacle.

Tables 1 and 2 below provide variable definitions and summary statistics, respectively.

Table 1. Variable description

Name	Definition	Source
<i>Expected Revenue Growth</i>		
REVINCR	A dummy variable of whether a firm expects its total sales to increase in the following year, coded 1 for “Increase” and 0 for “Decrease” or “Stay the same.”	BEEPS VI
<i>Perceived Ease of Access to Finance</i>		
FACCESS	Firm’s perception that access to finance is not an obstacle. It is coded as 0 for “Very severe obstacle”, 1 for “Major obstacle”, 2 for “Moderate obstacle”, 3 for “Minor obstacle” and 4 for “No obstacle”.	BEEPS VI
<i>Variables of firm climate change mitigation action</i>		
CO2MONITOR	Dummy =1 if the firm monitors its CO ₂ emissions; 0 otherwise.	BEEPS VI
ENERGYM	Dummy =1 if the firm adopts measures for energy management; 0 otherwise.	BEEPS VI
<i>Variable of female ownership</i>		
F_OWN	The percentage of the firm owned by females	BEEPS VI
<i>Variables for the firm top manager’s characteristics</i>		
F_TMAGER	Dummy=1 if the top manager is female; 0 otherwise.	BEEPS VI
TMANAGER_E XP	Logarithm of the top manager's years of experience in the sector.	BEEPS VI
<i>Control variables for firm characteristics</i>		
LNREVENUE	The natural logarithm of last year’s total sales	BEEPS VI
LNAGE	Logarithm of the firm’s age.	BEEPS VI
LISTED	Dummy = 1 if the firm is listed in a stock market; 0 otherwise.	BEEPS VI
MSME	Dummy = 1 if the firm is a micro, small, and medium-sized enterprise (MSME); 0 otherwise.	BEEPS VI
MANUF	Dummy = 1 if the firm is in the manufacturing sector; 0 otherwise.	BEEPS VI
WEATHER	Dummy = 1 if the firm experiences monetary losses from extreme weather events; 0 otherwise.	BEEPS VI
<i>Control variables for country</i>		
GG1Y, GG5Y, GG10Y	Average one-year, five-year, and ten-year real GDP growth of the country where the firm operates.	WBOD, IMF
INF1Y, INF5Y, INF10Y	Average five-year and ten-year inflation rates of the country where the firm operates.	WBOD
CORRUPT	The perception of the country's corruption. It is coded as 4 for “Very severe obstacle”, 3 for “Major obstacle”, 2 for “Moderate obstacle”, 1 for “Minor obstacle” and 0 for “No obstacle”.	BEEPS VI

4. Empirical results

4.1. The correlation between the firm’s climate actions and its expected revenue growth for the next year

Table 3 below presents the estimation results for the determinants of a firm’s expectation of increased total sales in the following year (REVINCR). Six specifications are reported, differing in the inclusion of short-term (one-year), medium-term (five-year), and long-term (ten-year) average macroeconomic controls, namely real GDP growth and inflation. All include area and year fixed effects and are weighted by population, with robust standard errors.

Table 2. Descriptive statistics of the variables

Variable	Obs	Mean	Std. Dev.	Min	Max
----------	-----	------	-----------	-----	-----

REVINCR	20,810	0.520	0.500	0.000	1.000
FACCESS	21,251	2.740	1.260	0.000	4.000
CO2MONITOR	21,091	0.040	0.200	0.000	1.000
ENERGYM	21,415	0.300	0.460	0.000	1.000
TMANAGER_EXP	21,198	20.850	11.540	1.000	70
F_TMAGER	21,702	0.160	0.360	0.000	1.000
F_OWN	21,208	15.610	31.260	0.000	100
LNREVENUE	19,516	12.990	3.050	-0.530	27.360
LNAGE	21,542	2.770	0.750	0.000	5.310
LISTED	21,567	0.070	0.260	0.000	1.000
MSME	21,698	0.780	0.410	0.000	1.000
MANUF	21,729	0.590	0.490	0.000	1.000
WEATHER	21,340	0.070	0.260	0.000	1.000
CORRUPT	20,010	1.420	1.410	0.000	4.000
GG1Y	21,729	2.520	3.500	-26.800	7.400
INF1Y	21,377	5.550	4.440	0.340	16.330
GG5Y	21,729	2.860	1.970	-1.620	6.860
INF5Y	21,729	6.370	6.380	-0.730	20.010
GG10Y	21,729	3.110	1.960	-0.980	8.830
INF10Y	21,729	6.190	5.020	0.600	21.580

The two climate change mitigation actions show strong positive associations with expected revenue growth. Specifically, firms that monitored their CO₂ emissions over the previous three years (CO2MONITOR) are approximately 7.9 percentage points more likely to expect an increase in sales in the coming year, holding all other variables constant. The effect is even larger for energy management measures (ENERGYM), which increases the probability of expecting revenue growth by about 11.4 percentage points. Both coefficients are statistically significant at the one percent level across all specifications, and their magnitude remains remarkably stable when controlling for different macroeconomic horizons. This suggests that proactive environmental practices are robustly correlated with managerial optimism regarding future sales.

Turning to gender-related leadership variables, the findings reveal contrasting patterns. A female top manager (F_TMAGER) is associated with a 4.7-percentage-point higher probability of expecting revenue growth. In contrast, the percentage of female ownership (F_OWN) exhibits a small but statistically significant negative coefficient of approximately -0.001 , implying that a ten-percentage-point increase in female ownership reduces the likelihood of expecting growth by only 0.1 percentage point, economically negligible, though precisely estimated.

Among firm-level control variables, several patterns emerge. Larger firms, as measured by log annual revenue (LNREVENUE), are more optimistic, with a coefficient of approximately 0.04. Conversely, older firms (LNAGE) are significantly less optimistic, reducing the likelihood of expecting growth by about 6 percentage points per unit increase in log age. Being publicly listed (LISTED) is weakly negatively associated with growth expectations (coefficient ≈ -0.029), while micro, small, and medium enterprises (MSMEs) are also less optimistic than large firms, with coefficients of -0.017 to -0.022 that are significant at conventional levels in most specifications. Manufacturing firms (MANUF) show a weak negative effect (-0.014 to -0.015). Having experienced monetary losses due to extreme weather events (WEATHER) is negatively associated with growth expectations only in models including ENERGYM (coefficient ≈ -0.038), but not in CO2MONITOR models, offering only partial support for the role of weather-related shocks.

Table 3. The correlation between the firm's climate actions and its expected increase in total sales for next year

	REVINCR Model 1	REVINCR Model 1	REVINCR Model 1	REVINCR Model 1	REVINCR Model 1	REVINCR Model 1
	OLS (1)	OLS (2)	OLS (3)	OLS (4)	OLS (5)	OLS (6)
CO2MONITOR	0.079*** (0.000)		0.078*** (0.000)		0.079*** (0.000)	
ENERGYM		0.114*** (0.000)		0.114*** (0.000)		0.114*** (0.000)
TMANAGER_EXP	-0.000 (0.831)	-0.000 (0.354)	-0.000 (0.881)	-0.000 (0.387)	-0.000 (0.878)	-0.000 (0.385)
F_TMAGER	0.047*** (0.002)	0.047*** (0.002)	0.047*** (0.002)	0.046*** (0.002)	0.047*** (0.002)	0.047*** (0.002)
F_OWN	-0.001*** (0.002)	-0.001*** (0.004)	-0.001*** (0.002)	-0.001*** (0.004)	-0.001*** (0.002)	-0.001*** (0.003)
LNREVENUE	0.040*** (0.000)	0.038*** (0.000)	0.040*** (0.000)	0.038*** (0.000)	0.040*** (0.000)	0.038*** (0.000)
LNAGE	-0.060*** (0.000)	-0.057*** (0.000)	-0.060*** (0.000)	-0.057*** (0.000)	-0.060*** (0.000)	-0.057*** (0.000)
LISTED	-0.029* (0.090)	-0.029* (0.079)	-0.029* (0.089)	-0.029* (0.080)	-0.028* (0.094)	-0.029* (0.085)
MSME	-0.022** (0.041)	-0.018* (0.096)	-0.021* (0.051)	-0.017 (0.116)	-0.021* (0.051)	-0.017 (0.115)
MANUF	-0.015* (0.057)	-0.015* (0.059)	-0.015* (0.063)	-0.014* (0.065)	-0.015* (0.063)	-0.014* (0.065)
WEATHER	-0.026 (0.183)	-0.039** (0.045)	-0.025 (0.202)	-0.038** (0.048)	-0.025 (0.198)	-0.038** (0.047)
CORRUPT	-0.019*** (0.000)	-0.018*** (0.000)	-0.019*** (0.000)	-0.019*** (0.000)	-0.019*** (0.000)	-0.019*** (0.000)
GG1Y	0.008 (0.223)	0.008 (0.220)				
INF1Y	-0.018** (0.028)	-0.017** (0.035)				
GG5Y			0.001 (0.986)	0.001 (0.969)		
INF5Y			-0.051** (0.015)	-0.049** (0.018)		
GG10Y					-0.060 (0.454)	-0.056 (0.477)
INF10Y					-0.091** (0.029)	-0.088** (0.034)
Constant	0.222** (0.016)	0.220** (0.016)	0.202* (0.053)	0.200* (0.053)	0.647 (0.134)	0.623 (0.146)
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Weighted by population	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16046	16201	16238	16397	16238	16397
Adjusted R ²	0.159	0.166	0.158	0.165	0.158	0.165

p-values in parentheses

* *p* < 0.1, ** *p* < 0.05, *** *p* < 0.01

Perceived corruption as an obstacle (CORRUPT) is among the most robust negative predictors. A one-unit increase in the corruption obstacle scale (where higher values indicate more severe obstacles) reduces the probability of expecting revenue growth by approximately 1.8 to 1.9 percentage points. This suggests that firms operating in more corrupt environments are systematically more pessimistic about future sales.

Macroeconomic controls exhibit an interesting pattern. Inflation averages, whether one-year,

five-year, or ten-year, are consistently negative and significant, with coefficients ranging from -0.017 (one-year) to -0.091 (ten-year). This indicates that firms in countries with higher past inflation are less likely to expect revenue growth, regardless of the time horizon considered. In contrast, real GDP growth rates (GG1Y, GG5Y, and GG10Y) are never statistically significant at conventional levels, although the ten-year coefficient is negative and marginally insignificant.

Table 4 below re-estimates the same six specifications as Table 3 but introduces an interaction term between female top manager (F_TMAGER) and perceived corruption (CORRUPT). This interaction reveals a nuanced conditional relationship that was masked in the simpler model. In Table 3, female top managers were unconditionally associated with a 4.7 percentage-point higher probability of expecting revenue growth. However, Table 4 demonstrates that this apparent advantage is entirely driven by firms operating in corrupt environments. When corruption is perceived as no obstacle, female-led firms show no significant difference from male-led firms. Conversely, as corruption becomes more severe, the positive effect of having a female top manager increases substantially. For a firm facing very severe corruption (CORRUPT = 4), a female top manager raises the likelihood of expecting sales growth by approximately 13.1 percentage points relative to a male-led firm. Moreover, corruption imposes a 2.2 percentage point penalty per unit increase on male-led firms, whereas female-led firms are essentially insulated from this penalty. We also conduct a significance test for the marginal effect. Table 5 below shows that female-led firms have an expected revenue growth of 60.4% on average across all levels of corruption. In other words, female-led firms outperform male-led firms by 4.6 percentage points on average.

The positive and significant interaction suggests that the advantage of female leadership for revenue expectations is not universal but emerges specifically in institutionally weak environments where corruption is a severe obstacle. To understand why, we draw on the institutional voids theory (Khanna and Palepu, 2010). When formal institutions fail, female leaders excluded from corrupt male networks develop resilient internal capabilities as substitutes. They invest in transparent accounting, formal compliance, process efficiency, and human capital, creating organizational architectures that operate predictably despite external corruption. These internal capabilities improve operational reliability and reduce costs, directly supporting revenue growth. Additionally, female executives build alternative informal networks, such as women's business associations, cross-sector alliances, and connections with impact investors, that provide access to non-traditional financing and information on corruption avoidance (Smallbone, 2006). This finding is also consistent with theories that female managers exhibit higher risk aversion and lower tolerance for unethical practices (Adams and Funk, 2012), leading them to avoid costly bribery and instead build formal institutional capabilities that prove resilient when corruption disrupts informal networks. Furthermore, selection effects may operate, and only the most capable female managers rise to top positions in corrupt settings (Ryan and Haslam, 2005). Consequently, female-led firms in highly corrupt environments are associated with greater optimism about future sales, whereas in clean environments no performance difference is observed.

Table 4. The correlation between the firm's climate actions and its expected increase in total sales for next year, adding the interaction between the female top manager and perceived corruption

	REVINCR OLS	REVINCR OLS	REVINCR OLS	REVINCR OLS	REVINCR OLS	REVINCR OLS
	Model 1 (1)	Model 1 (2)	Model 1 (3)	Model 1 (4)	Model 1 (5)	Model 1 (6)
CO2MONITOR	0.078*** (0.000)		0.077*** (0.000)		0.078*** (0.000)	
ENERGY		0.114*** (0.000)		0.114*** (0.000)		0.114*** (0.000)
TMAGER_EXP	-0.000 (0.826)	-0.000 (0.349)	-0.000 (0.876)	-0.000 (0.382)	-0.000 (0.873)	-0.000 (0.379)
F_TMAGER × CORRUPT	0.032*** (0.000)	0.034*** (0.000)	0.031*** (0.000)	0.034*** (0.000)	0.031*** (0.000)	0.034*** (0.000)
F_TMAGER	-0.001 (0.955)	-0.005 (0.804)	-0.001 (0.954)	-0.005 (0.807)	-0.001 (0.962)	-0.005 (0.813)
F_OWEN	-0.000*** (0.006)	-0.000** (0.010)	-0.001*** (0.005)	-0.000*** (0.009)	-0.001*** (0.005)	-0.000*** (0.009)
LNREVENUE	0.040*** (0.000)	0.038*** (0.000)	0.040*** (0.000)	0.038*** (0.000)	0.040*** (0.000)	0.038*** (0.000)
LNAGE	-0.059*** (0.000)	-0.057*** (0.000)	-0.060*** (0.000)	-0.057*** (0.000)	-0.060*** (0.000)	-0.057*** (0.000)
LISTED	-0.028* (0.094)	-0.029* (0.082)	-0.028* (0.092)	-0.029* (0.082)	-0.028* (0.098)	-0.028* (0.088)
MSME	-0.022** (0.042)	-0.018* (0.097)	-0.021* (0.052)	-0.016 (0.118)	-0.021* (0.051)	-0.017 (0.117)
MANUF	-0.016** (0.049)	-0.015** (0.048)	-0.015* (0.054)	-0.015* (0.054)	-0.015* (0.054)	-0.015* (0.054)
WEATHER	-0.026 (0.184)	-0.039** (0.044)	-0.025 (0.203)	-0.038** (0.048)	-0.025 (0.198)	-0.038** (0.046)
CORRUPT	-0.022*** (0.000)	-0.022*** (0.000)	-0.022*** (0.000)	-0.022*** (0.000)	-0.022*** (0.000)	-0.022*** (0.000)
GG1Y	0.008 (0.225)	0.008 (0.223)				
INF1Y	-0.018** (0.027)	-0.017** (0.033)				
GG5Y			0.000 (0.991)	0.001 (0.974)		
INF5Y			-0.051** (0.014)	-0.049** (0.017)		
GG10Y					-0.060 (0.448)	-0.057 (0.469)
INF10Y					-0.092** (0.028)	-0.088** (0.032)
Constant	0.229** (0.014)	0.227** (0.013)	0.208** (0.047)	0.206** (0.046)	0.658 (0.127)	0.636 (0.137)
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Weighted by population	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16046	16201	16238	16397	16238	16397
Adjusted R ²	0.160	0.167	0.159	0.166	0.159	0.166

p-values in parentheses

* *p* < 0.1, ** *p* < 0.05, *** *p* < 0.01

Table 5. Test of significance for the marginal effect

Dependent variable:

Delta-method

REVINCR	Margin	Std. Err.	t	P>t	[95% Conf. Interval]
F_TMAGER					
0 (Male)	0.558	0.004	139.84	0	0.550 0.565
1 (Female)	0.604	0.014	43.27	0	0.577 0.632

4.2. The correlation between the firm's climate actions and its perceived ease of accessing finance

Table 6 below presents the estimation results for the determinants of a firm's perceived ease of access to finance (FACCESS), with higher values indicating less difficulty obtaining financing. Six specifications are reported, differing in the inclusion of one-year, five-year, and ten-year average macroeconomic controls (real GDP growth and inflation). All included area and year fixed effects are weighted by population.

Turning first to climate change mitigation actions, a striking and unexpected pattern emerges. Firms that monitor their CO₂ emissions (CO2MONITOR) perceive significantly worse access to finance, with a coefficient of approximately -0.107 across all specifications. This implies that CO₂ monitoring is associated with a reduction in perceived financing ease of about one-tenth of a point on the four-point scale, holding other factors constant. The negative effect of CO₂ monitoring on financing perceptions may reflect that firms that monitor CO₂ emissions perceive worse access to finance. In contrast, the adoption of energy management measures (ENERGYM) shows no statistically significant association with FACCESS.

Regarding managerial characteristics, the top manager's industry experience (TMANAGER_EXP) exhibits a small but consistently positive and statistically significant effect across all columns (coefficient ≈ 0.002). Although the magnitude is modest, the robustness of this finding suggests that more experienced managers are better able to navigate financial markets, likely through accumulated networks, reputation, and knowledge of financing sources.

Larger firms, as measured by log annual revenue (LNREVENUE), perceive significantly easier access to finance, with a coefficient of approximately 0.020 . This aligns with the well-established literature showing that larger firms possess more collateral, formal financial documentation, and bargaining power with lenders. Micro, small, and medium enterprises (MSMEs) have worse access to finance than large firms, with coefficients of approximately -0.044 to -0.045 that are marginally significant. This finding is consistent with the extensive literature on SME financing constraints.

One of the most robust effects in Table 6 comes from weather-related shocks. Firms that experienced monetary losses due to extreme weather events (WEATHER) over the previous three years perceive substantially worse access to finance, with coefficients ranging from -0.156 to -0.166 . This reduction of approximately 0.16 points on the FACCESS scale is economically meaningful. The likely mechanism is that weather shocks damage firms' assets, reduce profitability, and increase firms' perceived riskiness from a lender's perspective.

Perceived corruption (CORRUPT) emerges as a negative predictor of access to financing. A one-unit increase in the corruption obstacle scale (where higher values indicate more severe obstacles) reduces perceived ease of access to finance by approximately 0.190 to 0.191 points. This suggests that corruption directly and severely impedes firms' ability to obtain financing. In corrupt environments, firms may face demands for bribes to secure loans, contract enforcement becomes unreliable, and lenders perceive higher country risk, all of which worsen perceptions of financing.

Finally, macroeconomic variables exhibit no statistically significant relationships with FACCESS. Neither one-year, five-year, nor ten-year averages of real GDP growth or inflation are statistically significant.

Table 6. The correlation between the firm's climate actions and its perceived ease of accessing finance

	FACCESS	FACCESS	FACCESS	FACCESS	FACCESS	FACCESS
	OLS	OLS	OLS	OLS	OLS	OLS
	Model 2	Model 2	Model 2	Model 2	Model 2	Model 2
	(1)	(2)	(3)	(4)	(5)	(6)
CO2MONITOR	-0.106** (0.035)		-0.107** (0.033)		-0.107** (0.032)	
ENERGYM		0.010 (0.657)		0.009 (0.684)		0.009 (0.681)
TMANAGER_EXP	0.002** (0.022)	0.002** (0.026)	0.002** (0.025)	0.002** (0.029)	0.002** (0.025)	0.002** (0.029)
F_TMAGER	0.042 (0.219)	0.041 (0.235)	0.042 (0.224)	0.040 (0.241)	0.042 (0.224)	0.040 (0.240)
F_OWN	-0.000 (0.671)	-0.000 (0.726)	-0.000 (0.665)	-0.000 (0.721)	-0.000 (0.666)	-0.000 (0.722)
LNREVENUE	0.020*** (0.000)	0.019*** (0.000)	0.020*** (0.000)	0.019*** (0.000)	0.020*** (0.000)	0.019*** (0.000)
LNAGE	-0.003 (0.830)	-0.003 (0.851)	-0.003 (0.824)	-0.003 (0.845)	-0.003 (0.815)	-0.003 (0.836)
LISTED	-0.034 (0.377)	-0.027 (0.482)	-0.034 (0.368)	-0.027 (0.470)	-0.034 (0.368)	-0.027 (0.469)
MSME	-0.044* (0.069)	-0.037 (0.119)	-0.045* (0.061)	-0.038 (0.105)	-0.045* (0.061)	-0.038 (0.105)
MANUF	-0.021 (0.230)	-0.025 (0.162)	-0.022 (0.221)	-0.025 (0.155)	-0.022 (0.219)	-0.025 (0.153)
WEATHER	-0.156*** (0.000)	-0.166*** (0.000)	-0.156*** (0.000)	-0.166*** (0.000)	-0.156*** (0.000)	-0.165*** (0.000)
CORRUPT	-0.190*** (0.000)	-0.191*** (0.000)	-0.190*** (0.000)	-0.191*** (0.000)	-0.190*** (0.000)	-0.191*** (0.000)
GG1Y	0.013 (0.395)	0.012 (0.438)				
INF1Y	0.013 (0.488)	0.012 (0.530)				
GG5Y			0.037 (0.675)	0.038 (0.665)		
INF5Y			-0.012 (0.790)	-0.016 (0.734)		
GG10Y					0.126 (0.483)	0.137 (0.443)
INF10Y					0.049 (0.603)	0.043 (0.646)
Constant	2.678*** (0.000)	2.701*** (0.000)	2.630*** (0.000)	2.646*** (0.000)	2.001** (0.040)	1.968** (0.043)
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Weighted by population	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16361	16524	16554	16721	16554	16721
Adjusted R ²	0.268	0.268	0.268	0.268	0.268	0.268

Table 7 below presents the interaction between female top managers and perceived corruption for perceived ease of access to finance (FACCESS). The interaction coefficient is negative and statistically significant (approximately -0.050), while the main effect of a female top manager becomes positive and significant (approximately 0.115). This implies that female-led firms enjoy substantially better perceived access to finance than male-led firms only in low-corruption environments. As corruption becomes more severe, this advantage diminishes and eventually reverses: at very high levels of corruption ($\text{CORRUPT} = 4$), female-led firms perceive their access as approximately 0.085 points worse than that of male-led firms. Moreover, corruption imposes a larger penalty on female-led firms: a one-unit increase in corruption reduces FACCESS by 0.235 points for female-led firms compared to 0.185 points for male-led firms. However, when we conduct a significance test for the marginal effect, the results show that female-led firms have a predicted FACCESS

score of 2.675 on average across all corruption levels. In other words, female-led firms have slightly higher perceived ease of accessing finance (+0.034) than male-led firms on average (see Table 8)

Table 7. The correlation between the firm's climate actions and its perceived ease of accessing finance, adding the interaction term between female top manager and perceived corruption

	FACCESS	FACCESS	FACCESS	FACCESS	FACCESS	FACCESS
	OLS	OLS	OLS	OLS	OLS	OLS
	Model 2	Model 2	Model 2	Model 2	Model 2	Model 2
	(1)	(2)	(3)	(4)	(5)	(6)
CO2MONITOR	-0.104** (0.038)		-0.105** (0.035)		-0.106** (0.034)	
ENERGY		0.010 (0.663)		0.009 (0.691)		0.009 (0.687)
TMAGER_EXP	0.002** (0.023)	0.002** (0.026)	0.002** (0.026)	0.002** (0.029)	0.002** (0.026)	0.002** (0.030)
F_TMAGER × CORRUPT	-0.047** (0.012)	-0.052** (0.005)	-0.047** (0.012)	-0.052** (0.005)	-0.047** (0.012)	-0.052** (0.005)
F_TMAGER	0.113** (0.011)	0.119** (0.007)	0.112** (0.011)	0.118** (0.007)	0.112** (0.011)	0.118** (0.007)
F_OWEN	-0.000 (0.534)	-0.000 (0.575)	-0.000 (0.527)	-0.000 (0.569)	-0.000 (0.528)	-0.000 (0.570)
LNREVENUE	0.021*** (0.000)	0.019*** (0.000)	0.020*** (0.000)	0.019*** (0.000)	0.021*** (0.000)	0.019*** (0.000)
LNAGE	-0.003 (0.821)	-0.003 (0.837)	-0.003 (0.814)	-0.003 (0.831)	-0.003 (0.806)	-0.003 (0.821)
LISTED	-0.034 (0.371)	-0.027 (0.476)	-0.035 (0.363)	-0.028 (0.464)	-0.035 (0.362)	-0.028 (0.464)
MSME	-0.044* (0.069)	-0.037 (0.119)	-0.045* (0.061)	-0.038 (0.105)	-0.045* (0.061)	-0.038 (0.105)
MANUF	-0.021 (0.245)	-0.024 (0.178)	-0.021 (0.235)	-0.024 (0.170)	-0.021 (0.234)	-0.024 (0.168)
WEATHER	-0.155*** (0.000)	-0.165*** (0.000)	-0.156*** (0.000)	-0.165*** (0.000)	-0.155*** (0.000)	-0.165*** (0.000)
CORRUPT	-0.185*** (0.000)	-0.185*** (0.000)	-0.185*** (0.000)	-0.185*** (0.000)	-0.185*** (0.000)	-0.185*** (0.000)
GG1Y	0.013 (0.391)	0.012 (0.434)				
INF1Y	0.013 (0.482)	0.012 (0.522)				
GG5Y			0.037 (0.675)	0.038 (0.666)		
INF5Y			-0.012 (0.801)	-0.015 (0.748)		
GG10Y					0.127 (0.480)	0.138 (0.440)
INF10Y					0.050 (0.594)	0.044 (0.635)
Constant	2.669*** (0.000)	2.690*** (0.000)	2.622*** (0.000)	2.636*** (0.000)	1.986** (0.041)	1.950** (0.044)
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Weighted by population	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16361	16524	16554	16721	16554	16721
Adjusted R ²	0.268	0.269	0.269	0.269	0.269	0.269

p-values in parentheses

* *p* < 0.1, ** *p* < 0.05, *** *p* < 0.01

Table 8. Test of significance for the marginal effect

Dependent variable: FACCESS	Delta-method		t	P>t	[95% Conf. Interval]
	Margin	Std. Err.			
F_TMAGER					
0	2.641	0.009	294.33	0	2.623 2.658
1	2.675	0.032	84.83	0	2.613 2.737

4.3. Robustness checks

To assess the robustness of our main findings, the baseline models were re-estimated using alternative econometric specifications. Specifically, a probit model was employed for expected revenue growth (REVINCR) in Model 1 because REVINCR has 2 values: 0 and 1. It is coded as 1 for “Increase” and 0 for “Decrease” or “Stay the same.” For Model 2, an ordered probit model was adopted for perceived ease of access to finance (FACCESS) because it is an ordinal variable ranging from 0 (“Very severe obstacle”) to 4 (“No obstacle”). All specifications included country fixed effects and country-level clustering.

Overall, the core results remain robust. As shown in Table 9, the two climate change mitigation actions, such as monitoring CO₂ emissions (CO2MONITOR) and adopting energy management measures (ENERGYM), continue to exhibit strong positive associations with expected revenue growth. When we introduce the interaction term between female top manager and perceived corruption (see Table 10), both CO2MONITOR and ENERGYM retain their positive and significant associations with expected revenue growth. However, the interaction term itself becomes statistically insignificant.

Turning to the relationship between climate actions and perceived access to finance (see Table 11), CO₂ emissions (CO2MONITOR) continue to show a significant negative association with perceived access to finance (FACCESS). Meanwhile, energy management measures (ENERGYM) show no statistically significant correlation with FACCESS. After including the interaction between female top manager and perceived corruption (see Table 12), the interaction term remains negatively associated with perceived access to finance. However, it does not materially change the main results.

Moreover, the variable CORRUPT is not mean-centered. For the robustness test, a new variable named CORRUPT_C was created to center it. Then we redo the analysis, and the key findings remain robust (see Tables 13 and 14).

Table 9. The robustness check for the correlation between the firm’s climate actions and its expected increase in total sales for next year, using the probit method

	REVINCR Model 1 probit (1)	REVINCR Model 1 probit (2)	REVINCR Model 1 probit (3)	REVINCR Model 1 probit (4)	REVINCR Model 1 probit (5)	REVINCR Model 1 probit (6)
CO2MONITOR	0.350*** (0.000)		0.349*** (0.000)		0.350*** (0.000)	
ENERGY		0.283*** (0.000)		0.283*** (0.000)		0.284*** (0.000)
TMAGER_EXP	-0.003* (0.073)	-0.004** (0.046)	-0.003* (0.099)	-0.003* (0.064)	-0.003* (0.098)	-0.003* (0.064)
F_TMAGER	0.035 (0.504)	0.045 (0.394)	0.033 (0.528)	0.043 (0.410)	0.033 (0.519)	0.044 (0.404)
F_OWEN	-0.001 (0.187)	-0.001 (0.139)	-0.001 (0.192)	-0.001 (0.140)	-0.001 (0.193)	-0.001 (0.140)
LNREVENUE	0.090*** (0.000)	0.085*** (0.000)	0.089*** (0.000)	0.083*** (0.000)	0.089*** (0.000)	0.084*** (0.000)
LNAGE	-0.137*** (0.000)	-0.136*** (0.000)	-0.138*** (0.000)	-0.137*** (0.000)	-0.137*** (0.000)	-0.137*** (0.000)
LISTED	-0.010	0.002	-0.006	0.007	-0.003	0.010

	(0.891)	(0.982)	(0.930)	(0.918)	(0.968)	(0.883)
MSME	-0.032	-0.023	-0.024	-0.017	-0.025	-0.018
	(0.402)	(0.550)	(0.525)	(0.670)	(0.501)	(0.646)
MANUF	-0.033	-0.031	-0.028	-0.026	-0.029	-0.026
	(0.518)	(0.550)	(0.570)	(0.608)	(0.565)	(0.603)
WEATHER	0.050	0.030	0.054	0.032	0.052	0.030
	(0.301)	(0.541)	(0.264)	(0.520)	(0.276)	(0.537)
CORRUPT	-0.020	-0.018	-0.021	-0.020	-0.021	-0.020
	(0.317)	(0.361)	(0.267)	(0.307)	(0.276)	(0.317)
GG1Y	0.011	0.009				
	(0.253)	(0.385)				
INF1Y	-0.046***	-0.048***				
	(0.001)	(0.001)				
GG5Y			-0.014	-0.009		
			(0.671)	(0.795)		
INF5Y			-0.109***	-0.108***		
			(0.001)	(0.002)		
GG10Y					-0.191	-0.170
					(0.155)	(0.225)
INF10Y					-0.219***	-0.222***
					(0.002)	(0.002)
Constant	-0.768***	-0.808***	-0.658***	-0.716***	0.064	-0.004
	(0.000)	(0.000)	(0.004)	(0.003)	(0.861)	(0.992)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (Country)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16046	16201	16238	16397	16238	16397
Adjusted R ²						

p-values in parentheses

* *p* < 0.1, ** *p* < 0.05, *** *p* < 0.01

Table 10. The robustness check for the correlation between the firm's climate actions and its expected increase in total sales for next year when adding the interaction between female top manager and perceived corruption, using the probit method

	REVINC Model 1 probit (1)	REVINC Model 1 probit (2)	REVINC Model 1 probit (3)	REVINC Model 1 probit (4)	REVINC Model 1 probit (5)	REVINC Model 1 probit (6)
CO2MONITOR	0.349*** (0.000)		0.349*** (0.000)		0.350*** (0.000)	
ENERGY		0.283*** (0.000)		0.283*** (0.000)		0.284*** (0.000)
TMAGER_EXP	-0.003* (0.077)	-0.004** (0.048)	-0.003 (0.103)	-0.003* (0.067)	-0.003 (0.102)	-0.003* (0.067)
F_TMAGER × CORRUPT	0.026 (0.439)	0.027 (0.415)	0.025 (0.452)	0.026 (0.432)	0.025 (0.448)	0.026 (0.428)
F_TMAGER	0.001 (0.976)	0.009 (0.820)	0.001 (0.988)	0.010 (0.814)	0.001 (0.981)	0.010 (0.811)
F_OWEN	-0.001 (0.185)	-0.001 (0.137)	-0.001 (0.189)	-0.001 (0.138)	-0.001 (0.191)	-0.001 (0.138)
LNREVENUE	0.090*** (0.000)	0.085*** (0.000)	0.088*** (0.000)	0.083*** (0.000)	0.089*** (0.000)	0.083*** (0.000)
LNAGE	-0.137*** (0.000)	-0.136*** (0.000)	-0.138*** (0.000)	-0.138*** (0.000)	-0.137*** (0.000)	-0.137*** (0.000)
LISTED	-0.009 (0.900)	0.002 (0.973)	-0.006 (0.938)	0.008 (0.911)	-0.002 (0.976)	0.011 (0.876)
MSME	-0.031 (0.401)	-0.024 (0.548)	-0.024 (0.523)	-0.017 (0.668)	-0.025 (0.500)	-0.018 (0.644)

MANUF	-0.033 (0.515)	-0.031 (0.547)	-0.028 (0.568)	-0.026 (0.605)	-0.029 (0.562)	-0.026 (0.600)
WEATHER	0.050 (0.301)	0.030 (0.541)	0.054 (0.264)	0.032 (0.520)	0.052 (0.276)	0.030 (0.537)
CORRUPT	-0.023 (0.303)	-0.022 (0.336)	-0.025 (0.262)	-0.024 (0.293)	-0.025 (0.269)	-0.024 (0.300)
GG1Y	0.011 (0.249)	0.009 (0.379)				
INF1Y	-0.046*** (0.001)	-0.048*** (0.001)				
GG5Y			-0.014 (0.680)	-0.009 (0.804)		
INF5Y			-0.110*** (0.001)	-0.108*** (0.002)		
GG10Y					-0.192 (0.151)	-0.171 (0.220)
INF10Y					-0.220*** (0.002)	-0.223*** (0.002)
Constant	-0.761*** (0.000)	-0.800*** (0.000)	-0.652*** (0.003)	-0.709*** (0.002)	0.077 (0.833)	0.010 (0.980)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (Country)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16046	16201	16238	16397	16238	16397
Adjusted R ²						

p-values in parentheses

* *p* < 0.1, ** *p* < 0.05, *** *p* < 0.01

Table 11. The robustness check for the correlation between the firm's climate actions and its perceived ease of accessing finance, using the ordered probit method

	FACCESS Model 2 Ordered probit (1)	FACCESS Model 2 Ordered probit (2)	FACCESS Model 2 Ordered probit (3)	FACCESS Model 2 Ordered probit (4)	FACCESS Model 2 Ordered probit (5)	FACCESS Model 2 Ordered probit (6)
CO2MONITOR	-0.110* (0.052)		-0.112** (0.047)		-0.113** (0.045)	
ENERGY		-0.022 (0.700)		-0.025 (0.657)		-0.025 (0.659)
TMAGER_EXP	0.002 (0.276)	0.002 (0.264)	0.001 (0.331)	0.002 (0.314)	0.001 (0.332)	0.002 (0.315)
F_TMAGER	0.037 (0.262)	0.038 (0.265)	0.035 (0.301)	0.035 (0.305)	0.034 (0.302)	0.035 (0.306)
F_OWEN	-0.000 (0.901)	-0.000 (0.916)	-0.000 (0.929)	-0.000 (0.949)	-0.000 (0.932)	-0.000 (0.951)
LNREVENUE	0.027*** (0.000)	0.027*** (0.000)	0.027*** (0.000)	0.027*** (0.000)	0.027*** (0.000)	0.027*** (0.000)
LNAGE	0.024 (0.343)	0.023 (0.360)	0.022 (0.360)	0.022 (0.377)	0.022 (0.366)	0.021 (0.383)
LISTED	-0.093 (0.202)	-0.087 (0.223)	-0.090 (0.214)	-0.085 (0.233)	-0.090 (0.209)	-0.085 (0.228)
MSME	-0.011 (0.632)	-0.006 (0.806)	-0.019 (0.418)	-0.015 (0.563)	-0.019 (0.423)	-0.015 (0.571)
MANUF	-0.115*** (0.000)	-0.118*** (0.000)	-0.119*** (0.000)	-0.121*** (0.000)	-0.119*** (0.000)	-0.121*** (0.000)
WEATHER	-0.181*** (0.000)	-0.190*** (0.000)	-0.182*** (0.000)	-0.189*** (0.000)	-0.181*** (0.000)	-0.189*** (0.000)
CORRUPT	-0.248*** (0.000)	-0.248*** (0.000)	-0.245*** (0.000)	-0.245*** (0.000)	-0.245*** (0.000)	-0.245*** (0.000)

GG1Y	-0.003 (0.673)	-0.003 (0.600)				
INF1Y	0.024** (0.038)	0.022* (0.054)				
GG5Y			-0.006 (0.807)	-0.004 (0.846)		
INF5Y			0.013 (0.338)	0.007 (0.596)		
GG10Y					0.063 (0.396)	0.077 (0.310)
INF10Y					0.075 (0.145)	0.069 (0.179)
cut1						
Constant	-1.665*** (0.000)	-1.661*** (0.000)	-1.723*** (0.000)	-1.717*** (0.000)	-1.414*** (0.000)	-1.398*** (0.000)
cut2						
Constant	-0.861*** (0.000)	-0.860*** (0.000)	-0.921*** (0.000)	-0.917*** (0.000)	-0.612** (0.015)	-0.599** (0.017)
cut3						
Constant	-0.070 (0.579)	-0.070 (0.592)	-0.131 (0.366)	-0.129 (0.387)	0.178 (0.460)	0.189 (0.428)
cut4						
Constant	0.497*** (0.000)	0.496*** (0.000)	0.437*** (0.003)	0.438*** (0.003)	0.746*** (0.004)	0.756*** (0.003)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (Country)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16361	16524	16554	16721	16554	16721
Adjusted R ²						

p-values in parentheses

* *p* < 0.1, ** *p* < 0.05, *** *p* < 0.01

Table 12. The robustness check for the correlation between the firm's climate actions and its perceived ease of accessing finance when adding the interaction between female top manager and perceived corruption, using the ordered probit method

	FACCESS Model 2 Ordered probit (1)	FACCESS Model 2 Ordered probit (2)	FACCESS Model 2 Ordered probit (3)	FACCESS Model 2 Ordered probit (4)	FACCESS Model 2 Ordered probit (5)	FACCESS Model 2 Ordered probit (6)
CO2MONITOR	-0.109* (0.053)		-0.111** (0.048)		-0.112** (0.046)	
ENERGY		-0.021 (0.707)		-0.024 (0.664)		-0.024 (0.667)
TMAGER_EXP	0.002 (0.284)	0.002 (0.273)	0.001 (0.341)	0.001 (0.324)	0.001 (0.342)	0.001 (0.326)
F_TMAGER × CORRUPT	-0.046** (0.011)	-0.052*** (0.004)	-0.047*** (0.010)	-0.053*** (0.003)	-0.047*** (0.009)	-0.053*** (0.003)
F_TMAGER	0.103** (0.013)	0.112*** (0.007)	0.101** (0.013)	0.110*** (0.007)	0.101** (0.013)	0.111*** (0.007)
F_OWEN	-0.000 (0.908)	-0.000 (0.927)	-0.000 (0.934)	-0.000 (0.959)	-0.000 (0.938)	-0.000 (0.960)
LNREVENUE	0.027*** (0.000)	0.027*** (0.000)	0.027*** (0.000)	0.027*** (0.000)	0.027*** (0.000)	0.027*** (0.000)
LNAGE	0.024	0.024	0.023	0.022	0.023	0.022

	(0.334)	(0.350)	(0.351)	(0.367)	(0.356)	(0.373)
LISTED	-0.094	-0.089	-0.091	-0.086	-0.091	-0.087
	(0.198)	(0.216)	(0.208)	(0.226)	(0.204)	(0.221)
MSME	-0.011	-0.006	-0.019	-0.015	-0.019	-0.014
	(0.640)	(0.817)	(0.426)	(0.573)	(0.431)	(0.581)
MANUF	-0.115***	-0.118***	-0.119***	-0.121***	-0.119***	-0.121***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
WEATHER	-0.181***	-0.190***	-0.182***	-0.189***	-0.181***	-0.188***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CORRUPT	-0.242***	-0.241***	-0.239***	-0.237***	-0.239***	-0.237***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
GG1Y	-0.003	-0.003				
	(0.669)	(0.597)				
INF1Y	0.024**	0.023**				
	(0.035)	(0.049)				
GG5Y			-0.007	-0.006		
			(0.776)	(0.810)		
INF5Y			0.013	0.008		
			(0.324)	(0.567)		
GG10Y					0.065	0.079
					(0.381)	(0.296)
INF10Y					0.078	0.072
					(0.130)	(0.158)
cut1						
Constant	-1.651***	-1.644***	-1.711***	-1.702***	-1.390***	-1.371***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
cut2						
Constant	-0.847***	-0.843***	-0.909***	-0.903***	-0.588**	-0.571**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.017)	(0.021)
cut3						
Constant	-0.056	-0.053	-0.119	-0.115	0.202	0.217
	(0.653)	(0.680)	(0.408)	(0.440)	(0.394)	(0.354)
cut4						
Constant	0.511***	0.513***	0.449***	0.452***	0.770***	0.784***
	(0.000)	(0.000)	(0.002)	(0.002)	(0.002)	(0.002)
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster (Country)	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16361	16524	16554	16721	16554	16721
Adjusted R ²						

p-values in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 13. The robustness check for the correlation between the firm's climate actions and its expected increase in total sales for next year when adding the interaction between female top manager and perceived corruption, using Mean-center corrupt (CORRUPT_C)

	REVINCR Model 1 OLS (1)	REVINCR Model 1 OLS (2)	REVINCR Model 1 OLS (3)	REVINCR Model 1 OLS (4)	REVINCR Model 1 OLS (5)	REVINCR Model 1 OLS (6)
CO2MONITOR	0.078*** (0.000)		0.077*** (0.000)		0.078*** (0.000)	
ENERGY		0.114*** (0.000)		0.114*** (0.000)		0.114*** (0.000)
TMAGER_EXP	-0.000 (0.826)	-0.000 (0.349)	-0.000 (0.876)	-0.000 (0.382)	-0.000 (0.873)	-0.000 (0.379)
F_TMAGER × CORRUPT_C	0.032*** (0.000)	0.034*** (0.000)	0.031*** (0.000)	0.034*** (0.000)	0.031*** (0.000)	0.034*** (0.000)
F TMAGER	0.044***	0.043***	0.044***	0.043***	0.044***	0.043***

	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
F_OWN	-0.000***	-0.000**	-0.001***	-0.000***	-0.001***	-0.000***
	(0.006)	(0.010)	(0.005)	(0.009)	(0.005)	(0.009)
LNREVENUE	0.040***	0.038***	0.040***	0.038***	0.040***	0.038***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LNAGE	-0.059***	-0.057***	-0.060***	-0.057***	-0.060***	-0.057***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LISTED	-0.028*	-0.029*	-0.028*	-0.029*	-0.028*	-0.028*
	(0.094)	(0.082)	(0.092)	(0.082)	(0.098)	(0.088)
MSME	-0.022**	-0.018*	-0.021*	-0.016	-0.021*	-0.017
	(0.042)	(0.097)	(0.052)	(0.118)	(0.051)	(0.117)
MANUF	-0.016**	-0.015**	-0.015*	-0.015*	-0.015*	-0.015*
	(0.049)	(0.048)	(0.054)	(0.054)	(0.054)	(0.054)
WEATHER	-0.026	-0.039**	-0.025	-0.038**	-0.025	-0.038**
	(0.184)	(0.044)	(0.203)	(0.048)	(0.198)	(0.046)
CORRUPT_C	-0.022***	-0.022***	-0.022***	-0.022***	-0.022***	-0.022***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
GG1Y	0.008	0.008				
	(0.225)	(0.223)				
INF1Y	-0.018**	-0.017**				
	(0.027)	(0.033)				
GG5Y			0.000	0.001		
			(0.991)	(0.974)		
INF5Y			-0.051**	-0.049**		
			(0.014)	(0.017)		
GG10Y					-0.060	-0.057
					(0.448)	(0.469)
INF10Y					-0.092**	-0.088**
					(0.028)	(0.032)
Constant	0.197**	0.195**	0.176*	0.175*	0.626	0.604
	(0.033)	(0.032)	(0.091)	(0.090)	(0.147)	(0.158)
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16046	16201	16238	16397	16238	16397
Adjusted R ²	0.160	0.167	0.159	0.166	0.159	0.166

p-values in parentheses

* *p* < 0.1, ** *p* < 0.05, *** *p* < 0.01

Table 14. The robustness check for the correlation between the firm's climate actions and its perceived ease of accessing finance when adding the interaction between female top manager and perceived corruption, using Mean-center corrupt (CORRUPT_C)

	FACCESS Model 2 OLS (1)	FACCESS Model 2 OLS (2)	FACCESS Model 2 OLS (3)	FACCESS Model 2 OLS (4)	FACCESS Model 2 OLS (5)	FACCESS Model 2 OLS (6)
CO2MONITOR	-0.104** (0.038)		-0.105** (0.035)		-0.106** (0.034)	
ENERGY		0.010 (0.663)		0.009 (0.691)		0.009 (0.687)
TMAGER_EXP	0.002** (0.023)	0.002** (0.026)	0.002** (0.026)	0.002** (0.029)	0.002** (0.026)	0.002** (0.030)
F_TMAGER × CORRUPT_C	-0.047** (0.012)	-0.052*** (0.005)	-0.047** (0.012)	-0.052*** (0.005)	-0.047** (0.012)	-0.052*** (0.005)
F_TMAGER	0.046 (0.184)	0.045 (0.192)	0.045 (0.189)	0.044 (0.197)	0.045 (0.189)	0.044 (0.196)
F_OWN	-0.000 (0.534)	-0.000 (0.575)	-0.000 (0.527)	-0.000 (0.569)	-0.000 (0.528)	-0.000 (0.570)
LNREVENUE	0.021***	0.019***	0.020***	0.019***	0.021***	0.019***

	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
LNAGE	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003
	(0.821)	(0.837)	(0.814)	(0.831)	(0.806)	(0.821)
LISTED	-0.034	-0.027	-0.035	-0.028	-0.035	-0.028
	(0.371)	(0.476)	(0.363)	(0.464)	(0.362)	(0.464)
MSME	-0.044*	-0.037	-0.045*	-0.038	-0.045*	-0.038
	(0.069)	(0.119)	(0.061)	(0.105)	(0.061)	(0.105)
MANUF	-0.021	-0.024	-0.021	-0.024	-0.021	-0.024
	(0.245)	(0.178)	(0.235)	(0.170)	(0.234)	(0.168)
WEATHER	-0.155***	-0.165***	-0.156***	-0.165***	-0.155***	-0.165***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
CORRUPT_C	-0.185***	-0.185***	-0.185***	-0.185***	-0.185***	-0.185***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
GG1Y	0.013	0.012				
	(0.391)	(0.434)				
INF1Y	0.013	0.012				
	(0.482)	(0.522)				
GG5Y			0.037	0.038		
			(0.675)	(0.666)		
INF5Y			-0.012	-0.015		
			(0.801)	(0.748)		
GG10Y					0.127	0.138
					(0.480)	(0.440)
INF10Y					0.050	0.044
					(0.594)	(0.635)
Constant	2.406***	2.427***	2.359***	2.374***	1.723*	1.687*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.077)	(0.082)
Area FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16361	16524	16554	16721	16554	16721
Adjusted R ²	0.268	0.269	0.269	0.269	0.269	0.269

p-values in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.4. Addressing endogeneity issues

These findings should be interpreted with appropriate caution regarding causality. At the same time, the empirical results presented above reveal robust associations among climate mitigation actions, firm characteristics, institutional factors, and the two outcomes of interest: expected revenue growth and perceived access to finance. Several sources of endogeneity may bias the estimates, including reverse causality, omitted variable bias, and measurement error. This section discusses these concerns and the strategies employed to mitigate them.

The most prominent endogeneity concern is reverse causality: rather than climate actions causing higher growth expectations, firms that are already more optimistic or growth-oriented may be more likely to adopt energy management measures or monitor their CO₂ emissions. Optimistic managers may have longer investment horizons, greater willingness to experiment with new practices, or stronger beliefs that environmental investments will pay off. Similarly, firms with easier access to finance may invest in climate actions, rather than in actions that improve perceptions of financing. The cross-sectional nature of the BEEPS VI data (firms surveyed once, with retrospective and forward-looking questions) precludes definitive causal ordering.

To partially address this concern, the dependent variable is forward-looking (expected revenue growth for the next year). On the other hand, independent variables reflect the company's status before and at the time of the survey, including climate action, firm characteristics, and perceptions of corruption. This temporal alignment reduces but does not eliminate reverse causality, as unobserved firm-level optimism could simultaneously drive both past adoption of climate measures and current

growth expectations.

Unobserved firm-level heterogeneity may confound the estimated relationships. For example, managerial quality or corporate culture could influence both the propensity to adopt climate actions and the likelihood of expecting revenue growth. A more competent or forward-looking management team might simultaneously implement energy management systems and generate optimistic sales forecasts. Similarly, firms with stronger social capital or better banking relationships might find it easier both to monitor emissions and to access finance, thereby producing a spurious negative correlation between CO₂ monitoring and perceptions of financing if the true driver is an unmeasured relationship quality. To mitigate omitted variable bias, the empirical models include a rich set of control variables: firm size (logged revenue), firm age, sector (manufacturing indicator), legal status (publicly listed indicator), MSME classification, and area and year fixed effects. Area fixed effects absorb time-invariant unobserved heterogeneity at the regional level (e.g., local business culture, regulatory enforcement quality). Year fixed effects absorb common macroeconomic shocks.

The perceived ease of accessing finance variable (FACCESS) is a four-point scale, which reduces some measurement variability but remains perceptual. Ideally, objective measures of credit access (e.g., loan approval rates, interest spreads, credit limits) would complement perceptual measures. The results should therefore be understood as reflecting perceived constraints, which are nonetheless important for firm behavior: firms act on their perceptions, regardless of objective reality.

5. Discussion and limitations

5.1. Discussion

This study set out to examine the determinants of two critical firm-level outcomes: expected revenue growth and perceived access to finance. Drawing on a large, cross-country sample of enterprises from the BEEPS VI survey, the analysis yields several novel findings that extend existing literature and open new avenues for research. The discussion below highlights the most interesting and unexpected patterns, organized by thematic area.

The climate action paradox: a divergence between growth expectations and perceptions of financing. One of the most striking findings is the asymmetric relationship between different climate mitigation actions and the two outcomes. Both CO₂ monitoring and energy management are strongly and positively associated with expected revenue growth; firms adopting these practices are more likely to anticipate sales increases. This result aligns with the resource-based view, which holds that environmental efficiency measures reduce costs, enhance reputation, and signal forward-looking management (Porter and Van Der Linde, 1995; Stefan and Paul, 2008). From managers' own perspectives, climate actions are not a drag on performance but rather a correlate of optimism.

Firms that actively monitor CO₂ emissions reported significantly higher perceived financing obstacles. This can be explained by the fact that when firms begin measuring their carbon footprint, they may uncover substantial exposure to climate-related risks: dependence on fossil-fuel inputs, energy-intensive production processes, vulnerability to carbon taxes, or supply chain disruptions from weather events. In the absence of mandatory disclosure, many firms may not monitor their emissions and therefore need not report or confront them. Once monitoring begins, however, the firms become aware of their vulnerabilities and may communicate this (directly or indirectly) to the lenders. Risk-averse commercial banks, particularly those in emerging markets with limited experience in climate risk assessment, may respond by tightening credit terms, demanding higher collateral, or rejecting loan applications. Thus, the negative association arises not because monitoring worsens financial outcomes, but because monitoring reveals pre-existing risks that lenders then price into their credit decisions. This mechanism is grounded in the literature on asymmetric information and

environmental risk (Barth et al., 2008; Chava, 2014). Banks are known to penalize firms with poor environmental profiles, but the act of measurement itself can be a disclosure event. For example, research on the EU Emissions Trading Scheme shows that firms that first reported their emissions experienced stock price declines, not because emissions changed, but because information was revealed (Griffin et al., 2017).

Female leadership is an asset for growth expectations and financing. Firms led by a female top manager are more likely to expect sales growth. This suggests that female managers may bring strategic perspectives, stakeholder engagement skills, or adaptive capabilities that foster optimism about future revenue. However, female leadership shows no significant association with perceived access to finance. Surprisingly, the percentage of female ownership exhibits a small but statistically significant negative association with growth expectations. This pattern implies that gender effects are nuanced: female managers (day-to-day decision-makers) positively influence growth expectations, whereas female ownership (strategic control and capital allocation) does not, or even slightly reverses, that effect. One interpretation is that female-owned firms may face disadvantages (discrimination, sectoral sorting, or constrained networks) that offset any positive effects of gender on optimism. In contrast, female managers operate within existing ownership structures that buffer such disadvantages.

Corruption as a dominant constraint. Among all predictors, perceived corruption consistently exhibits one of the largest effects. A one-unit increase on the corruption obstacle scale reduces the probability of expecting sales growth and lowers perceived access to financing. These effects are larger than those of firm size, age, or any macroeconomic variable. This finding reinforces a substantial literature documenting corruption's damaging effects on firm performance (Fisman and Svensson, 2007; Aidis et al., 2008) and extends this literature to forward-looking expectations and perceptions of financing. Notably, corruption's effect on access to financing is nearly 10 times as large (in standardized terms) as the effect of CO₂ monitoring. This implies that for firms struggling to access credit, reducing corruption may yield greater dividends than adopting climate practices.

The most novel and theoretically provocative finding emerges from the interaction analyses. When examining the conditional role of female top managers across different levels of perceived corruption, female-led firms outperformed their male-led counterparts in growth expectations and access to finance, particularly in corrupt environments.

5.2. Limitation

Due to the cross-sectional nature of the BEEPS VI data, all estimates reported in this paper are correlational. Causation cannot be established, nor is it claimed to be established. The models control for a rich set of firm- and country-level covariates, include area, country, and year fixed effects, and employ temporal ordering in which the independent variables (e.g., past adoption of climate measures, past weather shocks) precede the dependent variables in the survey question wording. However, these steps address only observable confounding and do not eliminate the risk of unobserved heterogeneity, reverse causality, or self-selection.

Furthermore, the BEEPS VI data were collected primarily between 2018 and 2020, both before and during the COVID-19 pandemic. While this context provides a valuable baseline for understanding structural relationships among climate actions, corruption, gender, and firm outcomes, the pandemic has since disrupted global GDP, supply chains, and financial markets. Macroeconomic conditions and firm expectations may have shifted. Therefore, the findings should be interpreted as reflecting the pre-pandemic and pandemic periods, and future research using post-COVID data is necessary to test whether the associations documented here remain stable or have been altered by the pandemic and subsequent recovery. The BEEPS VI survey remains the most recent publicly available

wave of the Business Environment and Enterprise Performance Survey at the time of this analysis, as the EBRD has not yet released subsequent waves. We acknowledge this as a limitation and encourage replication when newer waves become available.

Finally, future research using panel data such as multiple waves of BEEPS or linked administrative data could employ firm fixed effects to control for time-invariant unobserved heterogeneity. Quasi-experimental approaches, such as difference-in-differences around regulatory changes (e.g., mandatory emissions reporting or anti-corruption campaigns), could provide stronger causal evidence. Instrumental variables, such as exogenous weather shocks for climate action adoption or regional corruption scandals for perceived corruption, could also help address endogeneity. In the absence of such designs, the present study aims to document robust associations and generate hypotheses for future causal testing.

6. Conclusion and implications

This study examined the correlates of firms' expected revenue growth and perceived access to finance using a large, cross-country dataset of enterprises from the BEEPS VI survey, combined with macroeconomic data from the World Bank. Given the cross-sectional nature of the data, all findings should be interpreted as robust associations rather than causal effects. With this caution in mind, the analysis yields several novel patterns that offer suggestive implications for policies and business strategies.

First, climate change mitigation actions are strongly and positively associated with expected revenue growth. However, a critical divergence emerges in perceived access to finance, and energy management shows no association, whereas CO₂ monitoring is associated with worse perceived access to finance. In other words, not all climate-related actions are seen the same way by firms and managers. Energy management, likely perceived as an efficiency-enhancing practice, is correlated with optimism about future sales, whereas CO₂ monitoring, possibly because of its association with regulatory disclosure or compliance risk, is associated with greater reported financing difficulties. It is important to emphasize that these are perceptual measures; we cannot conclude that lenders actually penalize CO₂-monitoring firms, only that such firms report greater obstacles in accessing finance.

Second, female top managers are associated with higher growth expectations. Still, female ownership shows a small negative effect, and corruption consistently ranks among the largest perceived constraints on both outcomes. The conditional analysis reveals a striking pattern; female-led firms report higher growth expectations in corrupt environments but simultaneously report worse access to financing as corruption rises. This suggests that female managers may overcome financing disadvantages through alternative means such as internal efficiency, operational controls, or non-traditional funding sources.

Third, perceived corruption emerges as one of the largest constraints on both outcomes, with effects substantially larger than those of climate actions or most firm-level controls. This finding reinforces the institutional literature on corruption's damaging role in private-sector development and extends it to forward-looking expectations and perceptions of financing.

The findings of this study carry several suggestive implications, though these should be treated as hypothesis-generating rather than prescriptive. For firm managers, the positive association between climate change mitigation actions, particularly energy management, and expected revenue growth

suggests that investing in such practices is correlated with managerial optimism about future performance. Managers may consider energy management as part of operational strategy rather than mere compliance. However, the finding that CO₂ monitoring is associated with worse perceived access to finance presents a strategic puzzle. Firms that monitor emissions might benefit from proactively communicating their broader risk management strategies and efficiency gains to lenders, framing emissions data within a narrative of transition readiness rather than regulatory vulnerability. Furthermore, the strong negative association between weather shocks and perceptions of financing suggests that lenders might consider post-disaster credit guarantees or flexible restructuring options for affected firms, thereby preventing temporary weather events from becoming persistent financing obstacles.

For policymakers, corruption is one of the most perceived constraints on both growth expectations and access to financing. Anti-corruption reforms such as transparent loan approval processes, strengthened contract enforcement, and reduced bribery demands may yield greater dividends for private sector development than macroeconomic stabilization policies, at least in the institutional contexts covered by our sample. Regarding climate policy, the divergence between CO₂ monitoring and energy management suggests that not all climate mitigation actions are associated with perceptions of financing in the same way. Policymakers should consider designing disclosure frameworks that do not inadvertently discourage monitoring, potentially through regulatory safe harbors or certification schemes that verify credible transition plans. Subsidies or technical assistance for energy management are also justified, given their positive association with growth optimism.

Finally, the positive association between female leadership and growth expectations suggests untapped potential. Policies that reduce barriers to female managerial advancement, such as mentorship programs, leadership training, and anti-discrimination measures, may foster more optimistic and growth-oriented firms. The small negative coefficient on female ownership, while not a policy priority given its magnitude, warrants monitoring rather than immediate intervention.

Acknowledgement

This research is fully funded by the University of Finance – Marketing, Vietnam.

References

1. Adams, R. B. and Funk, P. (2012). Beyond the glass ceiling: does gender matter?. *Management Science*, 58(2), pp. 219-235. <https://doi.org/10.1287/mnsc.1110.1452>
2. Aghion, P., Angeletos, G. M., Banerjee, A. and Manova, K. (2010). Volatility and growth: Credit constraints and the composition of investment. *Journal of Monetary Economics*, 57(3), pp. 246-265. <https://doi.org/10.1016/j.jmoneco.2010.02.005>
3. Aidis, R., Estrin, S. and Mickiewicz, T. (2008). Institutions and entrepreneurship development in Russia: A comparative perspective. *Journal of Business Venturing*, 23(6), pp. 656-672. <https://doi.org/10.1016/j.jbusvent.2008.01.005>
4. Allen, F., Qian, J. and Qian, M. (2005). Law, finance, and economic growth in China. *Journal of*

- Financial Economics*, 77(1), pp. 57-116. <https://doi.org/10.1016/j.jfineco.2004.06.010>
5. Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), pp. 99-120. <https://doi.org/10.1177/014920639101700108>
 6. Barrett, S. and Graddy, K. (2000). Freedom, growth, and the environment. *Environment and Development Economics*, 5(4), pp. 433-456. <https://doi.org/10.1017/S1355770X00000267>
 7. Barth, J. R., Caprio, G. and Levine, R. (2008). *Rethinking bank regulation: Till angels govern*. Cambridge University Press, Cambridge, EA.
 8. Bernanke, B. S. (1983). Irreversibility, uncertainty, and cyclical investment. *The Quarterly Journal of Economics*, 98(1), pp. 85-106. <https://doi.org/10.2307/1885568>
 9. Berrone, P., Cruz, C. and Gómez-Mejía, L. R. (2012). Socioemotional wealth in family firms. *Family Business Review*, 25(3), pp. 258-279. <https://doi.org/10.1177/0894486511435355>
 10. Boone, C. and Hendriks, W. (2009). Top management team diversity and firm performance: Moderators of functional-background and locus-of-control diversity. *Management Science*, 55(2), pp. 165-180. <https://doi.org/10.1287/mnsc.1080.0899>
 11. Boso, N., Story, V. M. and Cadogan, J. W. (2013). Entrepreneurial orientation, market orientation, network ties, and performance: Study of entrepreneurial firms in a developing economy. *Journal of Business Venturing*, 28(6), pp. 708-727. <https://doi.org/10.1016/j.jbusvent.2013.04.001>
 12. Bowen, A., Cochrane, S. and Fankhauser, S. (2012). Climate change, adaptation and economic growth. *Climatic Change*, 113(2), pp. 95-106. <https://doi.org/10.1007/s10584-011-0346-8>
 13. Chava, S. (2014). Environmental externalities and cost of capital. *Management Science*, 60(9), pp. 2223-2247. <https://doi.org/10.1287/mnsc.2013.1863>
 14. Connelly, B. L., Certo, S. T., Ireland, R. D. and Reutzel, C. R. (2011). Signaling theory: A review and assessment. *Journal of Management*, 37(1), pp. 39-67. <https://doi.org/10.1177/0149206310388419>
 15. Croson, R. and Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic Literature*, 47(2), pp. 448-474. <https://doi.org/10.1257/jel.47.2.448>
 16. Dasgupta, S., Laplante, B., Wang, H. and Wheeler, D. (2002). Confronting the environmental Kuznets curve. *Journal of Economic Perspectives*, 16(1), pp. 147-168. <https://doi.org/10.1257/0895330027157>
 17. De Haas, R. and Popov, A. (2023). Finance and green growth. *The Economic Journal*, 133(650), pp. 637-668. <https://doi.org/10.1093/ej/ueac081>
 18. Dhaliwal, D. S., Li, O. Z., Tsang, A. and Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *The Accounting Review*, 86(1), pp. 59-100. <https://doi.org/10.2308/accr.00000005>
 19. Eccles, R. G. and Klimenko, S. (2019). The investor revolution. *Harvard Business Review*, 97(3), pp. 106-116.
 20. Fisher, I. (1930). *The theory of interest: Kelley and Millman*. New York: Macmillan. New York.
 21. Fisman, R. and Svensson, J. (2007). Are corruption and taxation really harmful to growth? Firm level evidence. *Journal of Development Economics*, 83(1), pp. 63-75. <https://doi.org/10.1016/j.jdeveco.2005.09.009>
 22. Flammer, C. (2021). Corporate green bonds. *Journal of Financial Economics*, 142(2), pp. 499-516. <https://doi.org/10.1016/j.jfineco.2021.01.010>
 23. Geroski, P. A. (2000). *The growth of firms in theory and in practice*. In N. Foss & V. Mahnke (Eds.), *Competence, governance, and entrepreneurship* (pp. 168-186). Oxford University Press, Walton Street, Oxford. <https://doi.org/10.1093/oso/9780198297178.003.0008>
 24. Glass, C., Cook, A. and Ingersoll, A. R. (2016). Do women leaders promote sustainability?

- Analyzing the effect of corporate governance composition on environmental performance. *Business Strategy and the Environment*, 25(7), pp. 495-511. <https://doi.org/10.1002/bse.1879>
25. Goyal, K. and Kumar, S. (2021). Financial literacy: A systematic review and bibliometric analysis. *International Journal of Consumer Studies*, 45(1), pp. 80-105. <https://doi.org/10.1111/ijcs.12605>
 26. Griffin, P.A., Lont, D.H. and Sun, E.Y. (2017). The Relevance to Investors of Greenhouse Gas Emission Disclosures. *Contemporary Accounting Research*, 34, pp. 1265-1297. <https://doi.org/10.1111/1911-3846.12298>
 27. Hallegatte, S. and Rozenberg, J. (2017). Climate change through a poverty lens. *Nature Climate Change*, 7(4), pp. 250-256. <https://doi.org/10.1038/nclimate3253>
 28. Hambrick, D. C. and Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review*, 9(2), pp. 193-206. <https://doi.org/10.5465/amr.1984.4277628>
 29. Hart, S. L. (1995). A natural-resource-based view of the firm. *Academy of Management Review*, 20(4), pp. 986-1014. <https://doi.org/10.5465/amr.1995.9512280033>
 30. Hoffman, A. J. (2007). Carbon strategies: How leading companies are reducing their climate change footprint. University of Michigan Press, Greene St, Ann Arbor, US. <https://doi.org/10.3998/mpub.306055>
 31. Ibarra, H. (1992). Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative Science Quarterly*, 37(3), pp. 422-447. <https://doi.org/10.2307/2393451>
 32. Khanna, T. and Palepu, K. G. (2010). Winning in emerging markets: A road map for strategy and execution. Harvard Business Press, Boston, Massachusetts.
 33. Kölbel, J. F., Busch, T. and Jancso, L. M. (2017). How media coverage of corporate social irresponsibility increases financial risk. *Strategic Management Journal*, 38(11), pp. 2266-2284. <https://doi.org/10.1002/smj.2647>
 34. Kolk, A. and Pinkse, J. (2005). Business responses to climate change: Identifying emergent strategies. *California Management Review*, 47(3), pp. 6-20. <https://doi.org/10.2307/41166304>
 35. Littlewood, D., Decelis, R., Hillenbrand, C. and Holt, D. (2018). Examining the drivers and outcomes of corporate commitment to climate change action in European high emitting industry. *Business Strategy and the Environment*, 27(8), pp. 1437-1449. <https://doi.org/10.1002/bse.2194>
 36. Malmendier, U. and Nagel, S. (2011). Depression babies: do macroeconomic experiences affect risk taking?. *The Quarterly Journal of Economics*, 126(1), pp. 373-416. <https://doi.org/10.1093/qje/qjq004>
 37. Malmendier, U. and Shen, L. S. (2024). Scarred consumption. *American Economic Journal: Macroeconomics*, 16(1), pp. 322-355. <https://doi.org/10.1257/mac.20210387>
 38. McGuire, G. M. (2002). Gender, race, and the shadow structure: A study of informal networks and inequality in a work organization. *Gender & Society*, 16(3), pp. 303-322. <https://doi.org/10.1177/0891243202016003003>
 39. Peters, G. F. and Romi, A. M. (2015). The association between sustainability governance characteristics and the assurance of corporate sustainability reports. *Auditing: A Journal of Practice & Theory*, 34(1), pp. 163-198. <https://doi.org/10.2308/ajpt-50849>
 40. Porter, M. E. and Van Der Linde, C. (1995). Toward a New Conception of the Environment-Competitiveness Relationship. *The Journal of Economic Perspectives*, 9(4), pp. 97-118.
 41. Post, C. and Byron, K. (2015). Women on boards and firm financial performance: A meta-analysis. *Academy of Management Journal*, 58(5), pp. 1546-1571. <https://doi.org/10.1257/jep.9.4.97>

42. Rajan, R. G. and Zingales, L. (1998). Financial Dependence and Growth. *The American Economic Review*, 88(3), pp. 559–586.
43. Ryan, M. K. and Haslam, S. A. (2005). The glass cliff: Evidence that women are over-represented in precarious leadership positions. *British Journal of Management*, 16(2), pp. 81-90. <https://doi.org/10.1111/j.1467-8551.2005.00433.x>
44. Smallbone, D. (2006). *Enterprising Women in Transition Economies* (F. Welter, Ed.) (1st ed.). Routledge, London, United Kingdom. <https://doi.org/10.4324/9781315256368>
45. Stefan, A. and Paul, L. (2008). Does it pay to be green? A systematic overview. *Academy of Management Perspectives*, 22(4), pp. 45-62. <https://doi.org/10.5465/amp.2008.35590353>



RESEARCH ARTICLE

Macprudential Policy and the Risk Exposure of Commercial Banks

Nafiseh Keshtgar*

Faculty of Literature and Humanities, University of Birjand, Birjand, Iran

Seyed Hossein Mirjalili

Institute for Humanities and Cultural Studies, Tehran, Iran

How to cite this article:

Keshtgar, N. and Mirjalili, S. H. (2026). Macprudential Policy and the Risk Exposure of Commercial Banks. Iranian Journal of Accounting, Auditing and Finance, 10(3), 79-93. doi: 10.22067/ijaaf.2026.47852.1599
https://ijaaf.um.ac.ir/article_47852.html

ARTICLE INFO

Article History

Received: 2025-10-08

Accepted: 2026-02-14

Published online: 2026-07-12

Keywords:

Banking Risk, Commercial Banks, Financial Stability, Macprudential Policy, Systemic General Method of Moments

Abstract

This paper evaluates the effectiveness of macroprudential policy in reducing commercial banks' risk exposure in Iran. In this study, an empirical model is developed using dynamic panel data techniques and the generalized method of moments (GMM) estimator to examine the impact of macroprudential policy on the risk-taking behavior of Iranian banks during the period 2010-2024. The analysis evaluates the potential effects of four macroprudential instruments aligned with policy objectives: the loan-to-income ratio (LTI), the loan-to-deposit ratio (LTD), the leverage ratio (LVR), and liquidity requirements (LIQ). As a result of stronger macroprudential oversight, commercial banks are significantly less likely to take on risk. In the context of financial sanctions and Iran's unique economic environment, macroprudential tools remain highly effective in mitigating banking risks. As such, LTD instruments have a time lag, whereas LVR instruments have an immediate impact on risk. Additionally, no single instrument is more effective than the composite macroprudential index at reducing risk-taking. A more secure, less vulnerable banking system in Iran is possible through macroprudential policies, as the results show.

<https://doi.org/10.22067/ijaaf.2026.47852.1599>



NUMBER OF REFERENCES
33



NUMBER OF FIGURES
1



NUMBER OF TABLES
5

Homepage: <https://ijaaf.um.ac.ir>
 E-ISSN: 2717-4131
 P-ISSN: 2588-6142

*Corresponding Author: Nafiseh Keshtgar
 Email: nkeshtgar@birjand.ac.ir
 Tel: 09156551774
 ORCID: 0009-0002-9938-6654

1. Introduction

The growing dangers commercial banks face in the current global financial landscape underscore the need for more robust risk management systems. The importance of macroprudential regulation in protecting financial institutions from both known and unknown threats has been widely acknowledged (Acharya et al., 2022). For countries like Iran, where financial sanctions, macroeconomic instability, and ongoing policy uncertainty all heighten credit and operational risks for banks, this strategy is crucial.

In recent decades, Iran's banking system has faced persistent challenges, including high levels of non-performing loans, insufficient capital, and weak regulatory oversight. In addition to disrupting economic growth, capital flows, imports, and exchange rates, financial sanctions have exacerbated these problems (Heydarian et al., 2021, 2022, 2023, 2025). As a result, commercial banks have become increasingly exposed to credit, operational, and liquidity risks and are frequently facing liquidity pressures. The role that banks play in financing investment and supporting macroeconomic growth makes them a direct threat to both economic and financial stability.

Considering the volatile exchange rates, economic uncertainty, and chronic inflation, Iranian banks must manage their risks effectively. There is a lack of data about the effectiveness of macroprudential policies in Iran, despite claims that they can stabilize banks and dampen financial cycles.

Despite the growing body of international evidence on the effectiveness of macroprudential policy, current knowledge remains highly concentrated in advanced economies or in cross-country panel settings. Iran's unique macro-financial context is characterized by persistent sanctions, chronic inflation, high exchange rate volatility, and structural fiscal imbalances (Safari et al., 2024).

Also, the Iranian studies that have been done so far mostly rely on aggregate financial data and don't go into enough detail about how different banks differ. Also, private banks in Iran differ significantly in size, ownership structure, funding models, and exposure to foreign currencies. This suggests that policy tools may not work the same way on all institutions. There isn't much evidence about these effects on distribution right now, so policymakers don't have clear information on how to choose tools and focus them.

Finally, the interaction between macroprudential actions and episodes of sanctions, as well as volatile macroeconomic conditions, remains unexplored. The topic of which of the capital-based tools, borrower-level restrictions, or liquidity requirements is more effective in containing risk in moments of strong external pressure is still an unresolved empirical subject. Filling this gap is critical to develop robust regulatory frameworks to protect financial stability in fragile emerging markets. With a focus on sanctions dynamics and institutional heterogeneity, this study demonstrates how macroprudential policy can influence Iranian commercial banks' risk exposure.

Accordingly, the main research question of this study is: "What impact does macroprudential policy have on the risk exposure of commercial banks in Iran?" and how can this effect be explained under the country's specific economic conditions?

In light of this, the paper examines the effectiveness of macroprudential policy in mitigating Iranian commercial banks' risk-taking behavior, using the Z-score as a measure of insolvency risk and of banks' ability to absorb unexpected losses. Over the period 2010–2024, we examined the impact of macroprudential instruments on banking risks using dynamic panel data from 16 Iranian commercial banks and the GMM-SYS estimator. Macroprudential instruments include loan-to-income ratios (LTI), loan-to-deposit ratios (LTD), leverage ratios (LVR), and liquidity requirements (LIQ). This paper contributes to the broader literature on macroprudential regulation by offering new insights into how this policy operates in highly uncertain and sanctions-constrained economic environments.

We extended the time horizon to a 15-year panel data set, substantially broader than the 8-year

periods used previously. A model designed around the disaggregation of macroprudential instruments into four transmission-channel categories. By integrating sanctions into the System GMM framework, the study isolates the net effect of macroprudential policy from the dominant impact of external shocks. With this contribution, we provide a critical policy insight for economies under persistent pressure and demonstrate a level of modeling precision that Iranian banks have never achieved.

The remainder of this paper is structured as follows. Section 2 provides the theoretical background. Section 3 reviews the relevant empirical literature on macroprudential policy and banking risks. The methodology, including data, variables, and econometric models, is discussed in Section 4. Section 5 reports and discusses the empirical findings. Finally, Section 6 concludes the paper.

2. Theoretical Background

The central bank can rely on macroprudential supervision and regulation to monitor banking risks. The critical point is the notion of negative externalities arising from disruptions or failures in financial institutions, such as commercial banks, financial markets, or financial instruments.

In addition to providing credit to the economy, macroprudential policy typically leverages banks. However, with the expansion of capital market activities and market-based financing, macroprudential policymakers must also monitor systemic risks arising outside the banking system and develop and implement policy responses to contain these risks.

Credit growth caps have been used by some countries when credit expansion has been high and systemic risks have risen rapidly.

As [Borio \(2014\)](#) emphasizes, instruments such as the leverage ratio are specifically designed to limit excessive procyclical lending and constrain credit risk before macroeconomic shocks materialize. [Claessens \(2015\)](#) indicated that tools directly targeting households and credit markets tend to be highly effective in reducing financial vulnerabilities.

Macroprudential policy instruments can influence bank risk-taking behavior. Liquidity tools aim to reduce systemic liquidity risk by ensuring that banks maintain sufficient liquid assets to meet obligations under stress conditions. Interbank and funding markets are stabilized by macroprudential policy through the enforcement of minimum liquidity standards. The leverage ratio is a macroprudential tool designed to enhance the resilience of the banking system by limiting banks' exposure. By restricting excessive leverage, it curbs risk-taking behaviors and contains the buildup of systemic vulnerabilities.

Empirical literature review

In the last ten years, there has been a proliferation of empirical research on macroprudential policy (MPP), with studies varying along three key dimensions: (i) the number of countries studied and the type of institutions they belong to, (ii) the methodology and data structure used to conduct the research, and (iii) the particular tools and results that were analyzed. Instead of analyzing studies sequentially, this section synthesizes earlier findings across these categories to illustrate what is established, where results diverge, and how the Iranian example fits into and differs from the literature.

International cross-country and regional evidence: A large body of cross-country and multi-country banking studies finds that macroprudential tools are, on average, effective in restraining bank risk-taking and mitigating systemic vulnerabilities. Using bank-level data from advanced and emerging economies, [Altunbas et al. \(2018\)](#) and [Andries et al. \(2017\)](#) show that tighter capital-based and borrower-based instruments are associated with lower bank risk and reduced contributions to

systemic risk. Researchers in this field usually use dynamic panel or fixed-effects frameworks to account for persistence in risk metrics and rely on large panels of banks across different jurisdictions. Although the size and timeliness of the effects vary across institutional settings, a common thread in this study is that capital-related instruments (such as leverage or capital buffers) yield larger and more immediate benefits than borrower-based measures.

Single-country evidence from emerging economies: Complementary insights emerge from country-specific studies that exploit richer institutional detail. [Zhang et al. \(2018\)](#) and [Zhao et al. \(2023\)](#) found that under macroprudential supervision, banks in China and Vietnam are less likely to take risks, and that this effect is amplified when multiple instruments are used simultaneously ([Do et al., 2025](#)). These publications focus on heterogeneity across banks and use longer bank-level panels than cross-country research. They indicate that the effectiveness of macroprudential instruments is conditional on the balance-sheet structure and funding mechanisms. At the same time, results are sensitive to model specification and to whether instruments are examined individually or through composite indices, highlighting the importance of instrument-level analysis.

Empirical research focusing on Iran remains relatively limited and methodologically constrained. There is a systematic problem with the literature that uses short data spans and a narrow set of instruments, which prevents the identification of lagged effects and countercyclical dynamics crucial to macroprudential policy ([Borio, 2014](#)). This leaves a clear empirical gap in the provision of a disaggregated, time-consistent assessment of macroprudential instruments and their dynamic effects within Iran's specific institutional context.

Analysing the differences and identifying the gaps: All things considered, there is clear evidence from previous research that supports three main points: first, that banks are less likely to take risks when using macroprudential tools; second, that the effectiveness of these tools varies depending on the type of instrument and the characteristics of the bank; and third, that the results vary greatly depending on the context. However, there are still significant gaps. A limited amount of empirical evidence exists on Iran (i), which uses sufficiently long panels to identify delayed and cyclical effects, (ii) compares macroprudential instruments within a unified empirical framework, and (iii) explicitly takes financial sanctions and constrained monetary policy autonomy into account.

Linking international evidence to Iran's institutional context: The international literature largely abstracts from persistent external shocks such as comprehensive financial sanctions, which directly affect funding conditions, risk perceptions, and policy transmission. For Iran, sanctions constitute a structural constraint rather than a temporary disturbance, suggesting that results from advanced and typical emerging economies may not be directly transferable. Despite limited domestic evidence, macroprudential effectiveness has not been systematically examined to determine whether and how it differs between sanctioned and non-sanctioned periods, nor has it been examined how institutional characteristics such as central bank independence and Islamic banking structure affect instrument performance.

Overall, the worldwide empirical literature provides strong evidence that macroprudential policies matter for bank risk-taking, but it gives only partial guidance for the Iranian instance. However, it only gives some advice for the Iranian case. A structured review of previous research shows that the main problems that remain to be solved are how well instruments work, how they change over longer periods, and how macroprudential policy interacts with structural restraints unique to each country. Because of these gaps, there needs to be a focused empirical study of Iran that is both grounded in and distinct from the current international and domestic literature.

3. Methodology

To examine the effectiveness of macroprudential policy on the risk-taking behavior of commercial

banks, a Systemic Generalized Method of Moments (GMM-SYS) regression model was employed to assess its impact on banks' risk-taking behavior. This model is particularly suitable for addressing potential endogeneity concerns, which often arise from the dynamic relationship between macroprudential policy actions and bank risk variables.

This approach aligns with previous empirical studies, such as [Akinci and Olmstead-Ramsay \(2018\)](#) and [Ely et al. \(2021\)](#), in capturing the complex interactions between regulatory policies and risk factors in the banking system. The model is formulated as follows:

$$Risk_{it} = \beta_0 + \beta_1 Risk_{i,t-1} + \beta_2 MAPP_{it} + \sum \beta^b BankSpecific_{it} + \sum \beta^M Macro_t + u_{it} \quad (1)$$

Where $i=1, \dots, N$ and $t=1, \dots, t$ denote banks and time periods, respectively. In equation (1), $Risk_{it}$ denotes the bank risk variable for bank i at time t , which is regressed on its own lag. $MAPP_{it}$ denotes macroprudential policy tools, $BankSpecific_{it}$ includes control variables such as bank size, net loans-to-total-assets ratio, income diversification, cost management efficiency, and lending index. $Macro_t$ includes macroeconomic variables such as GDP growth (LGDP), inflation rate (GNP Deflator), and a financial sanction dummy variable (FSAN), and u_{it} is the error term.

According to Equation 1, $\sum \beta^b$ denotes the coefficients of the bank-specific factors that serve as independent variables in the model. In contrast, $\sum \beta^M$ denotes the coefficients of the macroeconomic variables included as another set of independent variables in the specification.

Measurement of Bank Risk: The Z-score indicates the probability of bank insolvency. [Lepetit and Strobel \(2015\)](#) argue that the logarithm of the Z-score represents a measure of bankruptcy risk. [Houston et al. \(2010\)](#) and [Fang \(2016\)](#) also confirm the use of the inverse Z-score as a proxy for bank default probability the higher the inverse Z-score, the greater the bankruptcy risk.

In this paper, the Z-score is used to measure the dependent variable, bank risk, as it is a widely used risk indicator in multiple studies, including [Laeven and Levine \(2009\)](#) and [Lepetit and Strobel \(2015\)](#). The Z-score is calculated as the sum of return on assets (ROA) and the equity-to-asset ratio (E/A) divided by the standard deviation of ROA:

$$Z_{it} = \frac{ROA_{it} + \frac{E_{it}}{A_{it}}}{Sd.(ROA)_{it}} \quad (2)$$

Where ROA_{it} denotes return on assets for bank i at time t ; EA denotes the equity-to-asset ratio; and $Sd.(ROA)_{it}$ denotes the standard deviation of asset returns. Due to the high skewness of the Z-score distribution, the natural logarithm is typically used as a transformation. Profitability and capitalization are correlated with Z-scores, as are income volatility and instability. Consequently, a higher Z-score indicates lower risk and a more stable bank ([De Nicolo et al., 2010](#)).

According to Figure 1, from 2010 to 2024, Khavar Miane Bank, Pasargad Bank, and Karafarin Bank exhibit the highest financial stability and thus the lowest risk-taking. In contrast, Sarmayeh Bank, Ayande Bank, Shahr Bank, and Day Bank demonstrate the lowest financial stability and the highest risk-taking (Figure 1)¹.

¹ Data are extracted from banks' balance sheets.

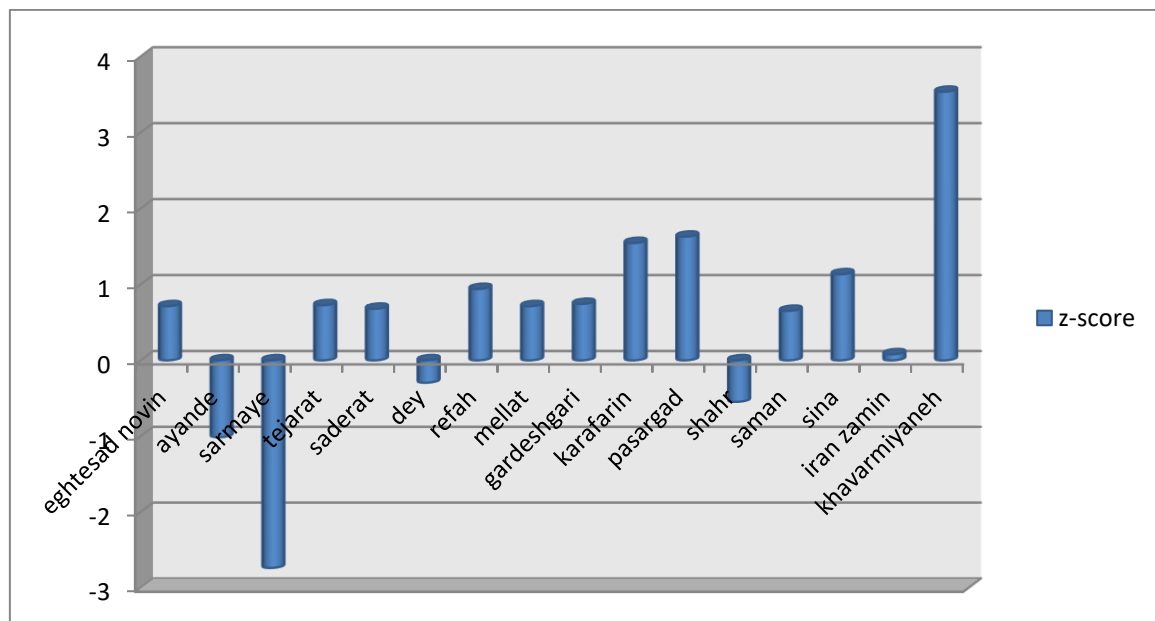


Figure 1. Average Z-score index for commercial banks during 2010–2024

Macroprudential Policy Tools: For conducting our analysis, we rely on the macroprudential policy indices available in the Integrated Macroprudential Policy database ¹ (iMaPP) developed by [Alam et al. \(2025\)](#), which is based on the annual review of macroprudential policy by the International Monetary Fund.

The macroprudential tools employed in the models are defined as follows. Among demand-based instruments, we selected the LTI ratio, which imposes limits on lending relative to borrower income. As a measure of capital relative to non-risk-weighted exposures, the LVR limits bank leverage and was utilized for capital requirements instruments. We use the LTD ratio, a tool from the loan-supply-based toolbox that regulates banks' lending relative to their deposit base and penalizes them for exceeding it. Finally, we implemented the Liquidity measure to meet the liquidity requirements. This measure seeks to reduce financing and systemic liquidity risks by establishing minimum standards for liquid asset ratios and liquidity coverage.

Before performing calculations, variables are normalized to ensure consistent units when determining the MPP index. We use the four weighted and normalized tools in accordance with the composite index method suggested by [Briguglio \(1995\)](#). Because different macroprudential tools target different dimensions of systemic risk, a composite macroprudential index is constructed by averaging them with equal weights. Moreover, using equal weights avoids subjective judgment about the supervisory priority of each instrument. The macroprudential literature provides no internationally accepted consensus on the relative importance of different instruments. This neutral, conservative approach helps avoid bias when evaluating the relative significance of the four instruments by assigning equal weights to each. The adoption of equal weights in this study is justified by the principles of neutrality and the absence of any theoretical or empirical basis for determining the relative importance of macroprudential instruments in the Iranian economy.

The normalization formula is defined as follows:

$$X_i^S = (X_i - \min X) / (\max X - \min X) \quad (3)$$

Where X_i denotes the observed value for period I , X_i^S denotes the standardized observation, and

¹ Integrated Macroprudential Policy Database

$\min X$ and $\max X$ are the minimum and maximum observed values, respectively. The MPP index is then calculated as:

$$LTI_t^S = (LTI_t - \min LTI) / (\max LTI - \min LTI) \quad (4)$$

$$LTD_t^S = (LTD_t - \min LTD) / (\max LTD - \min LTD) \quad (5)$$

$$LVR_t^S = (LVR_t - \min LVR) / (\max LVR - \min LVR) \quad (6)$$

$$LCR_t^S = (LCR_t - \min LCR) / (\max LCR - \min LCR) \quad (7)$$

$$MAPPINDEX_t = (LTI_t^S + LTD_t^S + LVR_t^S + LCR_t^S) / 4 \quad (8)$$

Bank-Specific Characteristics: Three characteristics are considered. Bank size (SIZE), measured by the logarithm of total bank assets. Net loan to total assets ratio (NLTA). Cost efficiency (CI) is calculated as the ratio of operating costs to total income. Lending variable, derived from the logarithm of total credit. Income diversification index (DIVL), calculated as:

$$DIVL = 1 - |RI - OI| / TI \quad (9)$$

RI denotes the bank's share of interest income, non-interest income by OI, and total income by TI. Interest income refers to income from lending, net of depositor shares, and non-interest income includes items such as foreign exchange gains, fees, and commissions.

The control variables for bank-specific variables were size, loan-to-asset ratio, equity ratio, non-interest income-to-total income ratio, and cost efficiency. Among these, bank size is a key determinant of risk-taking behavior. Franchise value theory suggests that higher franchise value, including large size, may limit a bank's incentives to take risks. The [Danisman and Demirel \(2019\)](#) study found that higher loan-to-asset ratios indicate greater credit risk. Under the Modigliani-Miller framework, holding more equity can reduce the risk and required risk premium to compensate investors for default costs ([Allen et al., 2011](#)). According to the fixed-leverage hypothesis, banks may be more willing to accept higher risk if leverage is reduced by increasing equity ([Godlewski, 2005](#)). Numerous subsequent empirical studies confirm that the use of risk-based capital ratios reduces bank risk ([Ariff, 2008](#)). Many empirical studies support the view that diversification reduces instability, as highlighted in [Rossi et al. \(2009\)](#). However, diversification can also be a source of banking risk due to several factors: increased volatility in non-interest income, higher fixed costs, lower regulatory capital requirements for non-lending activities, lack of managerial expertise, insufficient information about new product markets, exposure to new types of risk, and the fact that non-interest income can be riskier and more correlated with interest income ([Acharya et al., 2006](#)). Additionally, franchise value theory predicts that reduced management efficiency will encourage banks to take on more risk. [Williams \(2004\)](#) found that poorly managed banks tend to extend lower-quality loans. [Mwaurah \(2013\)](#) noted that managerial ability significantly influences banks' risk-taking decisions. For macroeconomic factors, economic growth (measured by GDP growth rate) and inflation (measured by the Consumer Price Index) are considered control variables. [Salas and Saurina \(2002\)](#) showed that during periods of economic growth, Spanish banks tend to expand lending activities to increase market share, often at the expense of credit quality. [Boyd and Champ \(2003\)](#) argue that inflation reduces the supply of credit available to businesses, affecting both the banking sector and the broader economy. Inflation also lowers the real return on assets and savings while simultaneously encouraging borrowing. This increase in personal and bank lending can, in turn, raise the probability of default. Banks may respond to this by setting higher interest rates, which can mitigate the impact on real returns but also increase the likelihood of default. A summary of the variables is presented in

Table 1.

Table 1. Summary of variables	
Variable	Measurement
INDEPENDENT VARIABLE (RISK)	
Zscore	$\text{Ln}((\text{ROA} + \text{EA}) / \text{Sd.ROA}) * (-1)$
MACROPRUDENTIAL POLICY INSTRUMENTS	
MaPPLVR	Lim et al. (2011)
MaPPLiq	Lim et al. (2011)
MaPPLTD	Lim et al. (2011)
MaPPLTI	Lim et al. (2011)
MPI(Macprudential Policy Index)	Briguglio composit index (1995)
BANK SPECIFIC FACTORS	
TA (Bank size)	Natural logarithm of total assets
CI (Cost management)	Operational cost/Total income
Credit (credit growth)	Natural logarithm of Credit
NPL	Non-performing loans to total gross loans
DivI (Income diversification)	Non-interest Income/ Total income
MACRO FACTORS	
GDP (economic growth)	Real GDP growth rate
Inflation	GDP Deflator
Financial Sanction	Dummy 1 for intensified financial sanctions and 0 for non

Source: The authors synthesize from previous studies

In Table 1, the dependent variable, the Z-score, is accompanied by the independent variables, including macroprudential policy instruments, the Briguglio composite index, bank-specific variables, and macroeconomic factors such as inflation and financial sanctions.

4. Empirical testing of the model and findings

In this study, five different models were estimated, and the empirical results are presented in Table 4. Before estimating the models, it is necessary to evaluate the statistical properties of the data, as shown in Table 2, and the stationarity tests, as presented in Table 3. The results of the Levin, Lin, and Chu unit root test indicate that all variables in the model are stationary. Furthermore, the modified Wald test for heteroskedasticity rejects the null hypothesis of homoscedasticity. Table 4 presents the results of the model estimation.

Table 2. Descriptive statistics of variables.

Dependent variable	Mean	Std.Dev	Min	Max	obs
Zscore	0.022	5.786	-42.469	9.970	240
Main regressors (MaPP)					
MaPPLiq	453.765	3293.800	-80.744	44900	240
MaPPLTD	7.672	9.130	0.164	76.091	240
MaPPLTI	12.144	8.423	0.068	83.333	240
MaPPLVR	7.021	8.697	-11.834	70.899	240
MPI	0.321	0.173	-0.022	1.227	240
Control variables					
TA	41.606	50.611	-29.240	418.678	240
CI	1.063	1.514	0.000	12.568	240
Credit	12.085	1.920	6.214	17.475	240
NPL	0.150	0.168	0.000	0.942	240
DIVI	0.575	0.404	0.000	3.406	240
RGDP	4.298	9.662	-4.299	36.255	240
Inflation	28.180	15.673	9.000	54.500	240
Sanction	0.800	0.4008	0.000	1.000	240

Table 2 presents the descriptive statistics, including the mean, standard deviation, minimum and maximum values, and the number of observations, for the banking risk variable and the independent and control variables used in the estimations.

Table 3. Model Variables: Stationary Test

Variable	Levin-Lin-Chu Test Statistic	Probability	Degree of stationarity	Difference or level
Zscore	-3.450	0.000	I(0)	Stationary at the level
MaPPLiq	-223.800	0.000	I(0)	Stationary at the level
MaPPLTD	-3.733	0.000	I(0)	Stationary at the level
MaPPLTI	-2.469	0.006	I(0)	Stationary at the level
MaPPLVR	-2.019	0.021	I(0)	Stationary at the level
MPI	-3.2102	0.000	I(0)	Stationary at the level
TA	-6.520	0.000	I(0)	Stationary at the level
CI	-1.980	0.023	I(0)	Stationary at the level
Credit	-42.220	0.000	I(0)	Stationary at the level
NPL	-11.890	0.000	I(0)	Stationary at the level
DIVL	-2.030	0.025	I(0)	Stationary at the level
RGDP	-37.660	0.000	I(0)	Stationary at the level
Inflation	-8.700	0.000	I(0)	Stationary at the level
Sanction	-8.623	0.000	I(0)	Stationary at the level

Table 3 presents the results of unit root tests for the banking risk, independent, and control variables, all of which are found to be stationary in levels.

Table 4. Results of the study

	Model 1	Model 2	Model 3	Model 4	Model 5
Zscore _{t-1}	1.152***	0.825***	0.255	1.113***	0.957***
MaPPLTI _t	0.081				
MaPPLTI _{t-1}	0.034				
MaPPLTD _t		-0.005			
MaPPLTD _{t-1}		0.105**			
MaPPLiq _t			-0.018		
MaPPLiq _{t-1}			0.218***		
MaPPLVR _t				0.076**	
MaPPLVR _{t-1}				0.003	
MaPPIIndex _t					2.143
MaPPIIndex _{t-1}					1.257**
TA	1.466	-0.005**	-0.859*	2.363*	-0.480
CI	-0.812	-0.543***	-0.737***	-0.299	-0.764***
DIVL	0.239*	1.771***	1.387**	1.143	-0.222
LnCredit	-1.216	-0.000***	-0.000***	-1.462*	0.623
NTLA	-2.036	-1.360	-0.315	-2.407***	-2.159
LnGDP	0.076	-0.119	-0.102	-0.770*	0.028
Deflator	-0.016*	-0.002	-0.011***	-0.002	0.003
SAN	-0.614	-0.473**	-0.584	-1.235***	-0.342***

*, **, *** statistical significance at 10%, 5% and 1%.

Statistical significance and direction (positive or negative) of the relationship between banking risk and the five models are summarized in Table 4.

The estimated γ coefficient in our dynamic model is statistically significant, indicating a stable, highly persistent process in the credit risk of Iranian banks. In addition to model specification tests, the Sargan test confirms the robustness of the estimation approach. Consequently, the estimate of γ

is statistically reliable. It reflects the persistence of bank credit risk in the Iranian banking system.

The estimated dynamic coefficient γ in the baseline model lies in the interval $0 < \gamma < 1$, with a value close to 1. This result reflects a high degree of persistence in the Z-Score indicator of bank stability. Accordingly, shocks affect bank risk gradually over several periods and are consistent with risk-taking behavior in the Iranian banking system.

Using the (LTI) as the macroprudential indicator in the first model, we examine bank risk-taking using banking-specific variables (log of total assets, income diversification index, loan-to-asset ratio, cost efficiency) as well as macroeconomic variables (GDP growth rate, inflation, financial sanctions).

According to the results, the coefficient on the lagged Z-score variable increases the Z-score, indicating that an increase in the previous period's Z-score leads to an increase in the current period. According to the coefficient of 1.152, Iranian commercial banks exhibit a self-reinforcing dynamic, whereby conditions in the previous period become more severe in the subsequent period. This finding underscores the importance of macroprudential intervention, as in the absence of such tools, banking risk cycles may become self-amplifying.

The term self-reinforcing dynamics refers to high persistence and the strong influence of past-period risk on current-period risk. This concept is directly related to the coefficient on the lagged dependent variable (γ) in the GMM model. When γ is positive and large (close to 1), this indicates self-reinforcing dynamics, meaning that credit risk in the previous period plays a substantial role in determining current credit risk. In the Iranian banking system, a high γ coefficient indicates that stability or instability (Z-score risk) is a persistent property.

Table 5. Model specification tests

	Model specification tests									
	Model 1		Model 2		Model 3		Model 4		Model 5	
	statistic	prob	statistic	prob	statistic	prob	statistic	prob	Statistic	prob
Arellano and Bond AR(1)	-1.750	0.080	-2.457	0.014	2.024	0.042	-3.227	0.001	-2.195	0.028
Serial autocorrelation test AR(2)	-0.401	0.688	-1.268	0.205	1.038	0.299	-0.990	0.320	-1.089	0.275
Sargan test	4.308	1.000	1.239	1.000	6.152	1.000	4.360	1.000	4.995	1.000
Wald test	1779.140	0.000	1970.610	0.000	4894.470	0.000	4568.770	0.000	7314.690	0.000

Finally, Table 5 reports the results of the model specification tests, highlighting the combined effect of macroprudential instruments on commercial banks' risk-taking behavior.

Because the coefficient for LTI is positive at 0.081 and the coefficient for the lagged LTI is positive at 0.034, the LTI limit lowers bank risk. However, neither of these terms is statistically significant. This might be because Iran's banks don't use the LTI tool effectively, and other risk factors, such as balance sheet gaps and bad loans, are more important.

As a result of inflationary pressures and structural shocks, LTI is insignificant in Iran.

The second model uses the (LTD) as the macroprudential indicator. The coefficient of the lagged Z-score variable is positive and significant, at 0.825. The positive coefficient of 0.105 on the lagged LTD indicates that bank risk decreases as the Z-score improves, indicating reduced risk-taking by banks.

The results demonstrate that the impact of LTD on the Z-score is positive but small in the current period, but becomes positive and substantial with a lag. However, in the long run, it improves financial stability by lowering liquidity demands and adjusting loan portfolios. This highlights the importance of a long-term perspective in evaluating the effectiveness of macroprudential tools in Iran.

In the third model, the LIQ is used as the macroprudential indicator. According to the results,

although the coefficient of the lagged Z-score is positive, it is not statistically significant. This indicates that, when controlling for macroprudential tools, macroeconomic conditions, and bank-fixed heterogeneity, there is no clear evidence of short-term stability in bank risk, and shocks and policy interventions play a more dominant role than the system's internal memory.

The effect of the LIQ variable on the Z-score is insignificant, but its lagged effect is positive and significant at 0.218. LIQ can mitigate systemic risk in two ways: first, by directly influencing credit growth, and second, by creating a liquidity buffer that can be used to reduce the severity of liquidity crises when needed. The findings indicate the delayed impact of this macroprudential tool.

In Iran's banking system, the effects of liquidity regulations are gradual and medium-term in nature. In the context of GMM, the medium-term horizon is derived from the persistence parameter (γ) and is approximately 3–4 years, i.e., the time needed for half of the MPP effect on the Z-score to dissipate. This implies that MPP in Iran must be implemented with a long-term outlook rather than with expectations of immediate results.

In the fourth model, we utilize the effective LVR as the macroprudential indicator. According to the results, the coefficient of the lagged Z-score variable increases the Z-score. Since the coefficient is at 1.113, it indicates that the risk of Iranian commercial banks exhibits a self-reinforcing dynamic.

The positive and significant coefficient of the leverage ratio, at 0.076 in the current period, indicates that higher bank capital immediately increases stability and reduces bankruptcy risk. However, the lagged effect of this variable, although positive, is not significant. This suggests that the effect of capital strengthening in Iran's banking system is mostly immediate and short-term and may lose its persistence in the medium term due to structural issues such as low asset quality, accumulated losses, and the accounting nature of capital increases. A higher leverage ratio in Iran is an immediate means of improving banking stability, but it is not sufficient to guarantee long-term risk sustainability.

In the fifth model, the combined effect of macroprudential tools (MAPP) on bank risk-taking is examined alongside other variables. The coefficient of the lagged Z-score is positive and significant at 0.957, thereby increasing the Z-score. The coefficient of MAPP is 2.142 and statistically insignificant, while the lagged MAPP coefficient is positive and significant at 1.257. This indicates that the combined use of macroprudential tools has the greatest impact on the Z-score and, consequently, on bank risk-taking compared to the other models.

In all five models, the three bank-specific variables consistently enhance financial stability and reduce bank risk-taking by enhancing the Z-score. As a result, the debt-to-asset ratio (LVR) has the greatest impact on the Z-score.

The coefficient on bank size (SIZE) is negative and significant in two of the five models. Larger banks do not necessarily exhibit lower risk (i.e., higher Z-scores). This may be due to greater risk-taking by large banks that anticipate government support. The effect of bank size in Iran is nonlinear and complex. In the Iranian economy, larger bank size does not automatically imply greater stability. This finding aligns with international studies in developing countries, which show that large size, without risk management and asset quality control, can negatively affect banking stability.

Two out of five models have a positive and statistically significant income diversification index (DIVL) coefficient, suggesting that diversifying income sources made the banking industry more stable. Depending on the state of the economy and the quality of the income, income diversification may or may not help banks mitigate risk. In all models, the net loan-to-asset ratio (NLTA) coefficient is positive but not statistically significant. This suggests that lending volume is not a reliable risk indicator, particularly in countries like Iran, where lending is either mandatory or asset quality is low. The actual risk effect depends more on loan quality and credit management.

The insignificance of NLTA reflects the failure of asset-growth restrictions to mitigate the risks

embedded in non-productive assets within Iranian banks.

Three of five models have a negative and significant cost-efficiency index (CI), suggesting that higher cost-to-income ratios, indicative of less efficiency, are associated with higher risk. This highlights that cost control and operational efficiency are important factors in banking stability. The lending volume index (LCRE) coefficient is negative and significant in two out of five models, indicating that high credit volumes in some banks are associated with increased risk, likely due to concentration in high-risk or low-quality loans. This is in line with NLTA's insignificance and demonstrates that distribution and quality of loans are more important than quantity alone.

One of the macroeconomic variables that has a negative and significant impact on most models is financial sanctions (FSAN). This means that sanctions increase bank risk by restricting access to foreign resources, reducing liquidity, and creating uncertainty. As a result, the Z-score goes down. As a result of banks' reliance on government policies, mandated lending, or structural economic volatility, GDP growth is insignificant across all models with mixed signs, indicating that economic growth does not directly or sustainably affect bank risk in Iran. The inflation variable (GDP deflator) is negative and significant in two models, indicating that high inflation reduces banking stability; real asset values decline, and credit costs rise, thereby increasing risk.

In Iran, bank credit risk is primarily influenced by structural shocks rather than cyclical factors, and by multicollinearity between GDP and sanctions.

5. Conclusion and policy recommendations

The empirical evidence confirms that MPP is vital for curbing excessive risk-taking in Iran. However, its effectiveness depends heavily on the specific tool selected and its associated time horizon. A key finding is that, when viewed from an instrument-hierarchy perspective, leverage ratios are the best immediate risk mitigation. In contrast, long-term liquidity requirements are necessary for medium-term stability. Second, regarding the lag factor, macroprudential interventions in Iran do not yield immediate results, except for capital-based ratios, which require a long-term policy outlook of 3 to 4 years to achieve a significant impact. Third, on structural obstacles, the dominance of financial sanctions and persistent inflation reduces the efficacy of cyclical tools, necessitating a shift toward resilience-buffer instruments. Fourth, on internal vs. external drivers, bank-specific factors, specifically operational efficiency and loan quality, are more predictive of stability than macroeconomic growth. The ultimate solution to banking stability in Iran is to strengthen capital through leverage limits and to apply liquidity tools to mitigate the deep-seated risks posed by inflation and external shocks.

It is recommended that the Central Bank of Iran adopt a tiered strategy for implementing macroprudential tools to enhance banking stability. In this regard, the priority would be immediate stability through the leverage ratio. To this end, the action needed is to designate the leverage ratio (LVR) as the primary, central anchor of the macroprudential framework. The timeline would be short-term. The rationale is that, unlike other tools, the leverage ratio has an immediate impact on reducing bankruptcy risk and increasing stability. In contrast to demand-based tools, it is an immediate capital buffer that does not suffer from implementation delays.

The second priority would be medium-term resilience through liquidity controls. To this end, the action needed is to implement and maintain strict loan-to-deposit ratios and liquidity requirements (LIQ). The timeline would be medium-term (3–4 years). The rationale is that these tools exhibit a lagged effect. Their benefits in reducing liquidity pressures and credit growth only become significant after the system memory adjusts. Policymakers must maintain these restrictions for at least 3–4 years to allow half of the policy effect to manifest in the Z-score.

The third priority would be sequencing and tool combination. To this end, the necessary action is

to shift from isolated tool use to a coordinated macroprudential package. In terms of sequencing, the Central Bank needs to begin with LVR to stabilize capital, followed by the simultaneous deployment of LTD and LIQ to manage credit expansion over the medium horizon. The focus should be on moving away from demand-based tools like Loan-to-Income (LTI) ratios, which are currently ineffective in Iran due to high inflation and structural shocks.

The fourth priority would be addressing internal bank efficiency. To this end, the necessary action is to enforce strict cost-efficiency (CI) and loan-quality mandates rather than focus on lending volume. The rationale is that high credit volumes are currently associated with increased risk due to low-quality or mandated lending. Efficiency and internal risk management are more influential for stability than bank size or simple volume restrictions.

References

1. Acharya, V. V., Bergant, K., Crosignani, M., Eisert, T. and McCann, F. (2022). The anatomy of the transmission of macroprudential policies. *The Journal of Finance*, 77(5), pp. 2533-2575. <https://doi.org/10.1111/jofi.13170>
2. Acharya, V. V., Hasan, I. and Saunders, A. (2006). Should banks be diversified? Evidence from individual bank loan portfolios. *The Journal of Business*, 79(3), pp. 1355-1412. <https://doi.org/10.1086/500679>
3. Akinci, O. and Olmstead-Rumsey, J. (2018). How effective are macroprudential policies? An empirical investigation. *Journal of Financial Intermediation*, 33, pp. 33-57. <https://doi.org/10.1016/j.jfi.2017.04.001>
4. Alam, Z., Alter, A., Eiseman, J., Gelos, G., Kang, H., Narita, M., Nier, E and Wang, N. (2025). Digging deeper—Evidence on the effects of macroprudential policies from a new database. *Journal of Money, Credit and Banking*, 57(5), pp. 1135-1166. <https://doi.org/10.1111/jmcb.13130>
5. Allen, F., Carletti, E. and Marquez, R. (2011). Credit market competition and capital regulation. *The Review of Financial Studies*, 24(4), pp. 983-1018. <https://doi.org/10.1093/rfs/hhp089>
6. Altunbas, Y., Binici, M. and Gambacorta, L. (2018). Macroprudential policy and bank risk. *Journal of International Money and Finance*, 81, pp. 203-220. <https://doi.org/10.1016/j.jimonfin.2017.11.012>
7. Andries, A. M., Melnic, F. and Nistor, S. (2017). Effects of macroprudential policy on systemic risk and bank risk taking. *Czech Journal of Economics and Finance, Forthcoming*, 68(3), pp.1-22. <https://dx.doi.org/10.2139/ssrn.3021278>
8. Ariff, M. and Ahmad, N. H. (2008). Multi-country study of bank credit risk determinants. *International Journal of Banking and Finance*, 5(1), pp. 135–152.
9. Borio, C. (2014). The financial cycle and macroeconomics: What have we learnt?. *Journal of Banking & Finance*, 45, pp. 182-198. <https://doi.org/10.1016/j.jbankfin.2013.07.031>
10. Boyd, J. and Champ, B. A. (2003). Inflation and financial market performance: what have we learned in the last ten years? *Monetary Policy in Low-Inflation Economies*, 1(1), pp. 259-301. <https://doi.org/10.26509/frbc-wp-200317>
11. Briguglio, L. (1995). Small island developing states and their economic vulnerabilities. *World Development*, 23(9), pp. 1615-1632. [https://doi.org/10.1016/0305-750X\(95\)00065-K](https://doi.org/10.1016/0305-750X(95)00065-K)
12. Claessens, S. (2015). An overview of macroprudential policy tools. *Annual Review of Financial Economics*, 7(1), pp. 397-422. <https://doi.org/10.1146/annurev-financial-111914-041807>
13. Danisman, G. O. and Demirel, P. (2019). Bank risk-taking in developed countries: The

- influence of market power and bank regulations. *Journal of International Financial Markets, Institutions and Money*, 59, pp. 202-217. <https://doi.org/10.1016/j.intfin.2018.12.007>
14. De Nicolo, G., Dell'Ariccia, G., Laeven, L. and Valencia, F. (2010). Monetary policy and bank risk taking. *IMF Staff Position Note*, New Hampshire, United States. <https://doi.org/10.2139/ssrn.1654582>
 15. Do, H. T., Pham, L. H., Ta, H. T. and Nguyen, H. T. D. (2025). Macroprudential policies and bank risk evidence from Vietnam. *Cogent Economics & Finance*, 13(1), pp. 1-13. <https://doi.org/10.1080/23322039.2025.2484653>
 16. Ely, R. A., Tabak, B. M. and Teixeira, A. M. (2021). The transmission mechanisms of macroprudential policies on bank risk. *Economic Modelling*, 94, pp. 598-630. <https://doi.org/10.1016/j.econmod.2020.02.005>
 17. Fang, Y. (2016). A study on the effectiveness of macro-prudential policies. *The Journal of World Economy*, 39(8), pp. 25-49.
 18. Godlewski, C. J. (2005). Bank capital and credit risk taking in emerging market economies. *Journal of Banking Regulation*, 6(2), pp. 128-145. <https://doi.org/10.1057/palgrave.jbr.2340187>
 19. Heydarian, S., Pahlavani, M. and Mirjalili, S. H. (2021). Financial sanctions and economic growth: an intervention time-series approach. *International Economics Studies*, 51(1), pp. 1-14. <https://doi.org/10.22108/IES.2020.122915.1083>
 20. Heydarian, S., Pahlavani, M. and Mirjalili, S. H. (2022). The impact of financial sanctions on capital Inflow and Outflow (case of Iran). *Journal of Money and Economy*, 17(1), pp. 67-88. <https://doi.org/10.29252/jme.17.1.67>
 21. Heydarian, S., Pahlavani, M. and Mirjalili, S. H. (2023). Financial sanctions and the imports of intermediate and capital goods in Iran: DID method. *International Journal of Business and Development Studies*, 15(2), pp. 101-134. <https://doi.org/10.22111/ijbds.2024.47819.2090>
 22. Heydarian, S., Pahlavani, M. and Mirjalili, S. H. (2025). Financial sanction, exchange rate volatility and macroeconomic variables (case of Iran), *Iranian Journal of Finance*, 9(2), pp.70-106. <https://doi.org/10.61186/ijf.2025.462663.1476>
 23. Houston, J. F., Lin, C., Lin, P. and Ma, Y. (2010). Creditor rights, information sharing, and bank risk taking. *Journal of Financial Economics*, 96(3), pp. 485-512. <https://doi.org/10.1016/j.jfineco.2010.02.008>
 24. Laeven, L. and Levine, R. (2009). Bank governance, regulation and risk taking. *Journal of Financial Economics*, 93(2), pp. 259-275. <https://doi.org/10.1016/j.jfineco.2008.09.003>
 25. Lepetit, L. and Strobel, F. (2015). Bank insolvency risk and Z-score measures: A refinement. *Finance Research Letters*, 13, pp. 214-224. <https://doi.org/10.1016/j.frl.2015.01.001>
 26. Lim, C. H., Columba, F., Costa, A., Kongsamut, P., Otani, A., Saiyid, M., Wezel, T and Wu, X. (2011). Macroprudential policy: what instruments and how to use them? Lessons from country experiences, *IMF Working Paper*, United States, pp. 1-59. <https://doi.org/10.2139/ssrn.1956385>
 27. Mwaurah, I. G. (2013). The Determinants of credit risk in commercial banks in Kenya (Doctoral dissertation), *University of Nairobi*, Nairobi, Kenya. <https://doi.org/10.21203/rs.3.rs-2038051/v1>
 28. Rossi, S. P., Schwaiger, M. S. and Winkler, G. (2009). How loan portfolio diversification affects risk, efficiency and capitalization: A managerial behavior model for Austrian banks. *Journal of Banking & Finance*, 33(12), pp. 2218-2226. <https://doi.org/10.1016/j.jbankfin.2009.05.022>
 29. Safari, O., Assadzadeh, A. and Mirjalili, S. H. (2024). Fiscal policy uncertainty and Industrial

- Investment in Iran. *International Journal of Business and Development Studies*, 16(1), pp. 229-259. <https://doi.org/10.22111/ijbds.2024.49134.2126>
30. Salas, V. and Saurina, J. (2002). Credit risk in two institutional regimes: Spanish commercial and savings banks. *Journal of Financial Services Research*, 22(3), pp. 203-224. <https://doi.org/10.1023/A:1019781109676>
31. Williams, J. (2004). Determining management behaviour in European banking. *Journal of Banking & Finance*, 28(10), pp. 2427-2460. <https://doi.org/10.1016/j.jbankfin.2003.09.010>
32. Zhang, X., Li, F., Li, Z. and Xu, Y. (2018). Macroprudential policy, credit cycle, and bank risk-taking. *Sustainability*, 10(10), pp. 1-18. <https://doi.org/10.3390/su10103620>
33. Zhao, Y., Goodell, J. W., Wang, Y. and Abedin, M. Z. (2023). Fintech, macroprudential policies and bank risk: Evidence from China. *International Review of Financial Analysis*, 87, pp. 1-28. <https://doi.org/10.1016/j.irfa.2023.102648>



RESEARCH ARTICLE

The Impact of ESG Performance on Capital Market Sensitivity

Behzad Kardan*, Hamideh Nazaridavaji

Department of Accounting, Faculty of Economics and Administrative Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

How to cite this article:

Kardan, B. and Nazaridavaji, H. (2026). The Impact of ESG Performance on Capital Market Sensitivity. *Iranian Journal of Accounting, Auditing and Finance*, 10(3), 95-111. doi: 10.22067/ijaaf.2026.48066.1621
https://ijaaf.um.ac.ir/article_48066.html

ARTICLE INFO

Article History

Received: 2025-10-03

Accepted: 2026-01-18

Published online: 2026-07-12

Abstract

This study examines the effect of environmental, social, and governance (ESG) performance on capital market sensitivity (CMS) among firms listed on the Tehran Stock Exchange. To assess how ESG performance influences firms' stock responses to market fluctuations, CMS was evaluated using three key indicators: beta, price volatility, and the Sharpe ratio. An analysis of 151 companies listed on the Tehran Stock Exchange from 2014 to 2023 was conducted in this quantitative, descriptive-analytical study. In addition to analyzing the direct effects of ESG performance on CMS, advanced econometric models were used to examine the moderating effects of inflation and economic sanctions. There was a statistically significant negative correlation between ESG performance and sensitivity to the capital markets. Stock price volatility was lower, and Sharpe ratios were greater for companies that performed well on ESG dimensions. In addition, the research on inflation and penalty intensity moderating effects showed that in bad political and economic circumstances, the impact of ESG performance on market sensitivity changes, either amplifying or diminishing, depending on the environmental factors. By concentrating on the Iranian capital market and incorporating the moderating effects of economic and political factors into the relationship between ESG performance and CMS, this study extended the existing literature on ESG and capital market risk management in emerging economies. Moreover, the study's applicability to investors and corporate managers made it more relevant.

Keywords:

ESG, CMS

<https://doi.org/10.22067/ijaaf.2026.48066.1621>



NUMBER OF REFERENCES
26



NUMBER OF FIGURES
0



NUMBER OF TABLES
8

Homepage: <https://ijaaf.um.ac.ir>
 E-ISSN: 2717-4131
 P-ISSN: 2588-6142

*Corresponding Author: Behzad Kardan
 Email: kardan@um.ac.ir
 Tel: 09153174057
 ORCID: 0000-0001-9205-3377

1. Introduction

The ESG framework has gradually evolved into a comprehensive framework for evaluating corporate performance across all ESG dimensions and has increasingly become a key criterion for investors (Zhao et al., 2018). This framework not only influences investment decisions but may also directly affect CMS. Sustainable and responsible investment opportunities are increasingly sought after by investors (Wu et al., 2022). According to Landi and Sciarelli (2019), companies with strong ESG performance are more likely to respond to market changes than others, driven by the growing demand for such investments. At the same time, firms that place greater emphasis on ESG-related issues tend to experience lower levels of risk, including ESG risks (Devalle et al., 2017). By allocating capital to such firms, investors may reduce the overall risk of their portfolios. This risk reduction, in turn, contributes to lower sensitivity of these firms' stock prices to market fluctuations (Sassen et al., 2016). Investors are increasingly favoring companies with strong ESG performance as ESG's importance grows. Investing behavior is expected to shift towards sustainable principles, resulting in higher valuations for companies. Accordingly, short-term volatility in stock prices may result from changes in investor expectations and increased competition for investments in ESG-oriented companies (Bermejo Climent et al., 2021).

In today's investment world, relying only on financial returns is no longer enough. Non-financial factors, such as ESG issues, have become important in investors' decision-making (Torricelli and Bertelli, 2022). In this situation, Iran is in a unique position as an economy in transition, facing many problems, including rising inflation and economic sanctions. The economic sanctions against Iran and the large swings in inflation have made it harder for businesses to operate and have affected investors' choices. Investor trust could be boosted if companies pay attention to ESG factors in these situations. Several contextual elements, such as business ownership structures, governmental involvement in the economy, and the degree of capital market growth, may influence the Iranian capital market's responsiveness to ESG variables. Consequently, the current study examined the influence of ESG performance on CMS in Iran. Considering Iran's specific economic conditions, including economic sanctions and high inflation, the need for research on the effect of ESG on CMS in this nation is especially evident. Since most research has focused on reputational risk (ESG) and financial performance (Sabouri and Allahyari, 2023), this study is notable. This research found that ESG reporting affects investors' risk perceptions (Bakhtiari, 2024), business performance risk (Aldieri et al., 2023), and investment efficiency (Hu et al., 2023). No study has, though, looked at this issue in the Iranian capital market.

Additionally, the study's main contribution is the examination of the moderating effects of economic sanctions and inflation on the correlation between ESG performance and sensitivity to capital markets. Thus, the study used these elements to investigate the impact of ESG on capital market responses to sanctions and extreme price volatility. This method may help Iranian politicians and investors make better judgments by making the results more applicable to their real-world situations. There is a lack of research in the Iranian capital market regarding the moderating effects of sanctions and inflation on the relationship between ESG and market sensitivity. Additionally, there is no investigation into the mechanisms through which ESG impacts market sensitivity in the context of sanctions and high inflation. Thus, the study contributes to the body of knowledge in this area by filling a gap in the literature. In other words, although both financial and non-financial factors have been shown to influence CMS in Iran, the present study demonstrates that, alongside financial activities, the environmental dimension of non-financial activities may also affect CMS in the Iranian context. Two other areas that could help improve future studies are refining methods for measuring CMS and developing ESG indices relevant to Iran's context. Concurrently, this study aids Iranian scholars, politicians, and market participants in the capital market by illuminating the effect of ESG

on the market in light of Iran's unique economic conditions, thereby enabling better decision-making. In today's complex financial markets, many factors affect asset values; it is crucial to identify and quantify them. Examining the relationship between ESG performance and CMS contributes to a clearer understanding of the role of non-financial factors in asset pricing. Investors, analysts, and lawmakers can all benefit from this kind of data when making decisions. This area of study also broadens understanding of investment and finance by posing new questions about how different ESG dimensions affect market sensitivity and interact with other economic variables. The study helps fill the gap in the Iranian capital market left by prior research. Furthermore, these kinds of studies not only contribute to what is already known but also lay the groundwork for future research. Overall, examining the relationship between ESG performance and CMS contributes to a more comprehensive understanding of financial market behavior and supports more informed investment decisions. With this information in hand, the boundaries of financial research can be pushed, and even more sophisticated financial products can be created. In light of the growing importance of ESG in investment decision-making and Iran's unique economic environment, analyzing the impact of ESG on corporate social responsibility is of great interest. Through a review of existing literature and an analysis of data collected from Iranian firms, this study attempts to shed light on the factors that moderate the impact of ESG on CMS in Iran. Therefore, the research questions addressed in the present study are as follows:

Is there a significant relationship between ESG factors and CMS among Iranian firms? Which factors influence the strength of this relationship? Do economic sanctions and inflation exert a moderating effect on the relationship between ESG performance and CMS?

2. Theoretical background and hypothesis development

ESG, as a central concept in the contemporary business environment, has increasingly influenced firm size and growth. Companies that follow ESG principles and perform better in ESG areas often have significant advantages over their competitors (Wu et al., 2022). Companies that follow ESG principles may be able to access new sources of funding, build stronger relationships with customers and business partners, and reduce the risks associated with climate change and social issues. In addition, companies that focus on ESG usually have a better image and credibility, which makes it easier to find and keep qualified workers. As a result, these companies are often able to grow faster and more sustainably than their competitors (Torricelli and Bertelli, 2022). On the flip side, if ESG principles are not adequately addressed, the firm's size can decrease. The Nasdaq (2019) reports that companies risk sanctions, heavy fines, and declining consumer demand if they damage the environment, violate workers' rights, or fail to maintain transparent corporate governance. A decrease in firm size, an increase in operational expenses, and a decrease in market share are all possible outcomes of these issues. According to Aydoğmuş et al. (2022), investors are increasingly investing in companies that adhere to ESG principles and avoiding those with poor ESG performance. One of the main factors influencing an organization's ownership structure is the firm's size. A smaller number of key shareholders, such as the founders, their families, or early-stage investors, tend to hold the majority of the shares in a smaller corporation. Among the reasons for this concentration of ownership are the need for seed money, the desire to keep management in charge, and the desire to save on monitoring costs. Owners of smaller businesses with larger responsibilities and a greater interest in maintaining the company's autonomy and character are more likely to be involved in management (Boshnak, 2025).

Companies' ownership arrangements also change as they grow bigger. Rahman and Hossain (2023) found that a wider variety of investors own a greater percentage of shares in large corporations.

Van Duuren et al. (2016) suggested that large enterprises are more attractive to institutional investors, and financial resources are needed to support growth and expansion. These companies often have institutional shareholders who want to get the most out of their money and have a lot of say in how the company runs. According to Akgun et al. (2021), managers' decision-making, information disclosure transparency, and investors' responses to changes can be influenced by the kind and concentration of ownership. The ownership structure of firms with a small number of large shareholders is less affected by market changes than that of firms with more dispersed ownership (Rahimzadeh et al., 2022). According to Zhou et al. (2022), these powerful shareholders tend to think in the long run and prioritize the development of long-term value. There is a greater likelihood that businesses with a dispersed ownership structure will be sensitive to market changes than companies with many small shareholders. Faster, more noticeable shareholder reactions to news and events are common in such organizations because no single shareholder has sufficient power to influence corporate decisions. The presence of institutional investors, such as pension funds and insurance companies, also plays an important role in shaping market sensitivity. The decisions of these investors are usually more influenced by their greater wealth and knowledge. They may place greater value on ESG factors because of their propensity for long-term, sustainable investments. Firm performance and information transparency can both be improved by institutional investors, thereby lessening the impact of market fluctuations (Bolibok, 2024). Shan et al. (2025) demonstrated that implementing ESG practices enhances financial stability and market attractiveness while also incurring additional operating costs. Investments in innovation and capital expenditures are also likely to rise among companies with better ESG performance. While investing in innovation can yield promising results, it is associated with more uncertain outcomes and has a smaller impact than capital expenditures, which greatly improve financial performance. Moreover, the financial performance of non-state-owned and non-polluting firms appears to be more sensitive to fluctuations in ESG performance. Similarly, Yang et al. (2025) reported a positive relationship between ESG performance and the use of an abnormally positive tone in annual reports. Mechanistic investigations also show that managers of companies that perform well in ESG areas are likely to use overly positive rhetoric, largely due to personal gains from equity-based pay. In an effort to alleviate the pressure from broader coverage by financial analysts and to enhance perceptions shaped by analysts' reports, managers of firms with higher ESG performance also tend to adopt a more aggressive tone in their annual disclosures.

Accordingly, in light of the foregoing discussion, a firm's ESG performance may directly influence its stock price. In addition, market regulators have increasingly recognized the importance of ESG considerations and have undertaken initiatives to encourage firms to disclose ESG-related information and to establish unified standards in this area. Such measures have contributed to greater market transparency and enhanced investor confidence (Bissoondoyal-Bheenick et al., 2023). In the end, ESG considerations have a significant role in how investors react to changes in the capital markets (Pu, 2023). There is less market volatility and more competitive advantage for firms that follow ESG principles and perform well in this area (Zhao et al., 2018). In light of the foregoing, the following is the formulation of the research hypothesis:

H1: ESG performance has a negative and statistically significant effect on CMS.

Moreover, consistent with the existing literature (Sabouri and Allahyari, 2023; Bakhtiari, 2024; Aldieri et al., 2023; Hu et al., 2023; Yang et al., 2025; Shan et al., 2025), the effect of ESG performance on CMS may be moderated by different environmental, economic, and political conditions. In particular, inflation and economic and political sanctions may alter the strength or direction of the ESG effect. Under conditions of high inflation, economic pressures, and rising operating costs, firms' resources and capacity to implement ESG strategies may be constrained (Bissoondoyal-Bheenick et al., 2023). In other words, inflation may constrain firms' financial

resources and pose challenges to implementing sustainable development projects, investments in green technologies, or social programs, thereby weakening the positive effect of ESG performance on reducing CMS.

On the other hand, economic sanctions and political pressures may restrict firms' access to international financial markets, foreign suppliers, and overseas investors (Pu, 2023). These constraints not only increase firms' operational and financial risks but also moderate the capacity of ESG performance to reduce volatility and enhance investor confidence. Simply put, when faced with sanctions-related pressures, even companies that excel in ESG metrics may not fully benefit from ESG's buffering impact on CMS. The importance of ESG factors is further highlighted by the fact that inflation and sanctions could increase economic uncertainty and systemic risk. Some protection against market volatility may be available to companies with strong ESG practices (high transparency, effective risk management, and sustainable performance). However, it will likely be less than under normal conditions. Hence, sanctions and inflation are moderating factors that might make ESG performance less effective at reducing market sensitivity or more effective. Therefore, when studying the relationship between ESG performance and CMS in emerging nations, it is important to carefully account for this moderating effect. In light of this, it is important to consider environmental and economic factors when studying the impact of ESG on market sensitivity; otherwise, our findings may be flawed. Inflation and sanctions can act as moderating factors, helping explain how ESG issues affect market stability and investor confidence.

3. Methodology

Using current financial data, this study employed a causal-correlational design and a quasi-experimental, ex post facto research methodology within the framework of positive accounting research. From the perspective of its nature and objectives, the research was applied, as applied studies aim to develop practical knowledge in a specific field. At the same time, with respect to data collection and analytical procedures, the study followed a causal–correlational approach.

3.1. Study population

In this research, all companies listed on the Tehran Stock Exchange between 2014 and 2023 were included in the statistical population.

3.2. Sampling method

Elimination sampling was employed systematically. After applying the following criteria, the final research sample was selected:

1. Firms had been listed on the Tehran Stock Exchange by the end of 2013.
2. Firms had provided complete financial information required for this study throughout the research period.
3. Firms were not classified as investment companies, banks, insurance companies, or other financial intermediaries.

Based on the information collected as of the end of 2023, the final sample was determined in accordance with Table 1.

3.3. Data collection methods and instruments

The required data for this study were collected from multiple sources, depending on their nature. Information related to the research literature and theoretical foundations was obtained from library-based sources, academic journals, and relevant online resources. Firm-level information, specifically balance sheets and income statements, was used as the primary research instrument.

Table 1. Number of firms in the statistical population and application of sample selection criteria

Description	Firms Excluded Over the Entire Period	Total Number of Firms
Total firms listed on the Tehran Stock Exchange		408
Exclusion of financial intermediation, financing, insurance, and investment firms	88	
Exclusion of firms that entered the stock exchange during the study period	55	
Exclusion of firms with trading suspension exceeding three months (due to the use of stock price volatility variables)	144	
Statistical population		151

Tadbir Pardaz¹, Rahavard Novin,² and other databases linked to the Tehran Stock Exchange, as well as reports issued by the Tehran Stock Exchange Organization, provided the raw data needed to test the research hypotheses. These data were obtained through direct reference to the sources, primarily by reviewing disclosure reports available on the CODAL platform, and were subsequently collected manually. In addition, comprehensive discs from the Tehran Stock Exchange Organization and rdis³ website were consulted, along with other relevant sources as necessary.

3.4. Data analysis procedure

The data analysis was conducted using a panel data structure, combining cross-sectional and annual observations. To test the research hypotheses, multivariate linear regression techniques were employed. Descriptive and inferential statistical methods were applied to analyze the collected data. The descriptive analyses were performed using frequency distribution tables. In contrast, the inference analyses were conducted using the F-Limer test, the Hausman test, normality tests, and multiple linear regression.

3.5. Research model

To test the research hypothesis, Model (1) was employed as follows:

Model (1)

$$MS_{it} = a_0 + a_1ESG_{it} + a_2Over.Con_{it} + a_3lev_{it} + a_4GRW_{it} + a_5size_{it} + a_6age_{it} + a_7roa_{it} + a_8loss_{it} + a_9BSIZE_{it} + a_{10}CFO_{it} + a_{11}TA_{it} + a_{12}Inflation_{it} + a_{13}Sanction_{it} + a_{14}Year_{it} + a_{15}Industry_{it} + \varepsilon_{it}$$

3.6. Research variables

3.6.1. Independent variable

ESG: Environmental, social, and governance reporting

The weighted framework employed in this study to measure ESG performance was adopted from Fakhari et al. (2018) and is specified as follows:

$$ESG = 0.337E + 0.388s + 0.275G$$

In this study, ESG represents the composite score of ESG reporting (according to Table 2). E denotes the environmental dimension score, S represents the social dimension score, and G refers to the governance dimension score.

1 <https://etadbir.portal.trade/>

2 <https://rahavard365.com/>

3 rdis.ir

Table 2. Indicators of environmental, social, and governance dimensions

Governance Dimension Indicators	Social Dimension Indicators	Environmental Dimension Indicators
Protection of shareholders' rights	Health, safety, and productivity	Pollution, emissions, and waste
Committees	Labor rights	Energy, efficiency, and water
Risk management	Stakeholder social perception	Biodiversity
Performance-based compensation	Product health and safety	Climate change
Ethics and codes of conduct	Non-discrimination and social inclusion	Environmental risk assessment
Corporate governance disclosure	Customer and supply chain social issues	Service and supply chain environmental issues
Regulatory compliance	Compliance with human rights principles	Environmental training
Board composition	Social development and philanthropy	Environmental management systems
Investment risk management	Corporate social responsibility investment	Environmental disclosure
Related-party transactions	Brand protection and anti-competitive behavior	—
Number of board meetings	Social learning and training	—
Corporate website availability	Social disclosure	—
Information timeliness	—	—
Ownership structure	—	—
Future plans	—	—
Employment reporting	—	—

After reviewing the board of directors' reports and the accompanying explanatory notes, a value of one was assigned if any of the indicators or their subcomponents were disclosed. In the final stage, each of the three dimensions was scored separately, and the aggregate ESG score was calculated using the weighted scheme $ESG = 0.337E + 0.388S + 0.275G$, through which the final ESG score was obtained.

3.6.2. Dependent variable

CMS (MS)

Beta: Beta is one of the most widely used indicators for measuring a stock's sensitivity to market movements. Compared to the market, a stock with a higher beta is more volatile.

Volatility: Volatility reflects the degree of price fluctuations of a stock or an index over a specific period. Higher volatility indicates greater market sensitivity.

Sharpe Ratio: An index of excess return relative to risk, derived from the standard deviation. When volatility increases while returns remain constant, the Sharpe ratio declines, indicating higher market sensitivity and is calculated as follows:

$$\text{Sharpe Ratio} = \frac{R_i - R_f}{\sigma_i}$$

where:

- R_i = The average annual return of the company i 's stock
- R_f = the risk-free rate (in this study, the rate of Iranian government bonds was used)
- σ_i = The standard deviation of the company i 's stock returns over the study period

3.6.3. Moderating variables

Sanction Intensity: The number of imposed sanctions, including financial, banking, oil-related, and other types, measures the severity of sanctions. Sanction intensity represents the level of

economic pressure and international restrictions imposed on a company or country, measured quantitatively. The sanction intensity index is derived based on the number and severity of sanctions imposed by recognized international and governmental institutions. The numerical scale of this index is normalized to range from 0 to 1, where 0 indicates no sanctions, and 1 represents the highest level of sanctions.

Inflation: In this study, inflation is measured using the annual Consumer Price Index (CPI). This index reflects the general price level of goods and services consumed by households in each year and is used as a continuous variable in the model. Using the annual CPI allows for the assessment of the impact of overall price levels on the study variables.

In Regression Model 2, the inflation and sanctions variables were entered as continuous variables to accurately capture their quantitative effects on the dependent variable. Using dummy variables for these factors could have led to excessive simplification and the loss of valuable quantitative information, thereby reducing the precision of the estimated effects. Therefore, including these variables as continuous variables was deemed more appropriate given the research objectives and the nature of the data. This approach allows examination of actual effects and small variations in variable values, resulting in more reliable and meaningful analytical outcomes.

3.6.4. Control variables

Managerial overconfidence (Over.CON): This study used an index based on excessive investment in assets to quantify managerial overconfidence. This metric follows the work of [Schrand and Zechman \(2012\)](#) by accounting for out-of-the-ordinary investments in assets, and is calculated as the residuals from regressing total asset growth on sales growth, with each industry-year's estimates assessed independently. The indicator is set to 1 if the regression residual is greater than 0, and to 0 otherwise. The use of this measure is grounded in the premise that, in firms where asset growth exceeds sales growth, managers tend to invest more aggressively than their peers, reflecting higher levels of managerial overconfidence.

$$\text{Assets.Gr}_{it} = a_0 + a_1 \text{sales.Gr}_{it} + \varepsilon_{it}$$

LEV (Financial Leverage): Leverage is the ratio of total liabilities to total assets for a given year.

Age (Firm Age): Firm age is defined as the length of time elapsed from the firm's establishment to the year under examination.

Size (Firm Size): Firm size is measured as the natural logarithm of total assets.

ROA (Return on Assets): ROA is calculated as net income divided by total assets for the current year.

GRW (Sales Growth): Sales growth is measured as the difference between current-year sales and prior-year sales, divided by prior-year sales.

Loss (Loss Indicator): During the past year, this variable equals 1 if the firm reported a loss and 0 otherwise.

TA (Asset Turnover): A firm's asset turnover is calculated by dividing its sales by its total assets.

Bsize (Board Size): Board size is defined as the number of board members in the year under review.

CFO (Operating Cash Flow): Operating cash flow is measured as operating cash flow divided by total assets.

Year: Year dummy variable.

Industry: Industry dummy variable.

4. Data analysis

4.1. Descriptive statistics

A qualitative analysis of CMS was conducted using data from 151 companies over the period 2014–2023. According to Table 3, the mean values of the dependent variables indicate that beta, with an average of 0.764, exhibits the highest mean among the sensitivity indicators, followed by stock price volatility (VOL) with a mean of 0.592 and the Sharpe ratio (SHARP) with a mean of -0.989 . These findings suggest that, within the sample, variations in firms' systematic risk (beta) are more pronounced than those reflected in the other indicators. At the same time, risk-adjusted performance, as measured by the Sharpe ratio, is relatively lower. A metric of stock price volatility shows 1.847 standard deviations, followed by the Sharpe ratio at 1.780, and beta at 0.774. This trend shows that companies' sensitivity to changes in the capital market varies widely. The independent variables indicate moderate dispersion and notable differences among firms in terms of ESG. A standard deviation of 1.236 is associated with the mean ESG performance score of 2.785. Mean values and acceptable levels of dispersion are also displayed by other control variables, such as board size (BSIZE), leverage (LEV), sales growth (GRW), and profitability (ROA). As the skewness and kurtosis statistics indicate, some variables have asymmetric distributions with relatively high values; for example, board size and stock price volatility. We should consider robust regression methods in future analyses. We find that indicators such as ESG_INFLATION and ESG_SANCTION have positive mean values and considerable dispersion in our descriptive statistical analysis of the interaction variables and market sensitivity measures. Thus, the effects of macroeconomic conditions and sanctions on firms' ESG performance are heterogeneous. This means the data are not normally distributed because both the Jarque-Bera statistics and their p-values are zero for every variable. Because of this, multiple regression models are well-suited to studying how ESG performance affects CMS, given the dataset's high dispersion and variability.

Table 3. Descriptive statistics of quantitative research variables

Variable	Mean	Median	Maximum	Minimum	Std. Dev.	Unit
BETA	0.764	0.765	4.180	-2.340	0.774	Index
ESG	2.784	2.625	8.539	1.100	1.235	Score
AGE	1.601	1.638	1.857	1.079	0.151	Log (years)
BSIZE	5.013	5.000	7.000	0.000	0.288	Number
CFO	0.124	0.102	0.726	-0.460	0.146	Ratio
GRW	0.471	0.372	18.169	-0.909	0.739	Percentage
INFLATION	0.330	0.412	0.523	0.090	0.164	Percentage
LEV	0.527	0.523	1.725	0.031	0.216	Ratio
LOSS	0.072	0.000	1.000	0.000	0.260	Dummy (0/1)
OVERCON	0.287	0.000	1.000	0.000	0.452	Dummy (0/1)
ROA	0.170	0.139	0.699	-0.400	0.161	Percentage
SANCTION	3.602	3.688	4.787	2.079	0.742	Index
SIZE	6.671	6.557	9.510	4.827	0.761	Log (assets)
TA	0.259	0.218	0.932	0.006	0.185	Ratio
ESG_INFLATION	0.915	0.790	3.242	0.099	0.631	Interaction
ESG_SANCTION	10.062	9.070	40.880	2.287	5.096	Interaction
VOL	0.592	0.075	24.234	-0.993	1.847	Percentage
SHARP	-0.988	-0.760	0.710	-45.550	1.780	Ratio

4.2. Multicollinearity test

To examine multicollinearity among the explanatory variables in the models, the variance inflation factor (VIF) test was used. All models' multicollinearity test results are identical because the explanatory variables are identical. Based on the calculated statistics, the VIF values for all variables

and the overall mean VIF across the regressions are below the 5 threshold. Consequently, there is no evidence of multicollinearity among the explanatory variables, indicating that multicollinearity does not pose a concern in the specified regression models.

4.3. Correlation matrix results

Pearson correlation coefficients were employed to examine the relationships among the variables included in the models. This test assesses the pairwise associations between the variables, and its results are reported in the correlation matrix presented in Table 4. The diagonal elements of the matrix, which represent the correlation of each variable with itself, are equal to one, indicating perfect correlation. Stronger positive correlations are observed near 1, whereas weaker correlations are observed near 0. Negative values indicate inverse relationships between variables. Correlation coefficients were positive and statistically significant at the 99% confidence level, indicating that the variables were related.

4.4. Estimation of the research model

Table 5 presents the results of the diagnostic tests for the first research model. The F-test, Hausman test, Breusch-Pagan test, serial correlation examination, and model specification test are all examples of these tests.

To assess the feasibility of pooling the data, the F-Limer test was applied. The null hypothesis of this test assumes that the data are poolable. The test statistic equals 1.276, which is greater than the critical F value. Accordingly, the null hypothesis is rejected, and it is concluded that the model should be estimated using panel data techniques.

Subsequently, the Hausman test was employed to determine whether a fixed-effects or random-effects specification is more appropriate. The test statistic for the Beta model is reported as 382.41, and the results for the other indicators are also statistically significant. Given that these values exceed the critical χ^2 values, the null hypothesis of random effects is rejected, indicating that the fixed-effects approach is the most suitable estimation method.

The Breusch-Pagan test was used to determine whether the regression residuals are homoskedastic. The results show that the test statistic is 92.672, which exceeds the critical χ^2 value. Therefore, the null hypothesis of homoskedasticity is rejected, implying the presence of heteroskedasticity in the regression residuals.

Based on the diagnostic test results, the research model was estimated using the fixed-effects approach. Table 6 reports the estimation results for the first model. CMS is negatively impacted by ESG performance, and the effect is statistically significant. Specifically, for the beta (Beta) indicator, the coefficient on ESG is -0.252 with a p-value of 0.0000, implying that a one-unit increase in ESG performance is associated with a 0.252-unit reduction in beta. There is a statistically significant difference between firms with higher ESG performance and firms with lower systematic risk. Regarding the volatility of stock prices (Vol), the ESG coefficient is -0.0647 , with a p-value of 0.0218. This indicates a strong correlation between stronger ESG performance and lower volatility in business returns, which in turn leads to more stable capital markets. For the Sharpe ratio (Sharpe), the ESG coefficient is -0.0811 with a p-value of 0.0218. This negative and statistically significant effect indicates that higher ESG performance is associated with a decline in the risk-adjusted return ratio. Rather than a decrease in absolute returns, this reduction is attributed to lower risk and market volatility. Some financial and economic variables, on the other hand, have distinct impacts. Sanctions (SANCTION) raise CMS, as shown by their positive and statistically significant impact on all indices (p-value = 0.0000).

Table 4. Result of diagnostic test of the first model

	Beta	Sharpe	Vol	ESG	AEG	BSIZE	CFO	GRW	INFLATION	LEV	LOSS	OVERCON	ROA	SANCTION	SIZE	TA
Beta	1.000															
Sharpe	0.004	1.000														
Vol	0.572	0.001	1.000													
ESG	0.191	0.018	0.002	1.000												
AEG	0.000	0.012	0.002	0.024	1.000											
BSIZE	0.003	0.623	0.015	0.030	0.010	1.000										
CFO	0.098	0.017	0.061	0.034	0.040	0.002	1.000									
GRW	0.016	0.053	0.008	0.178	0.081	0.071	1.540	1.000								
INFLATION	7.580	0.316	0.648	0.073	0.097	0.022	1.170	0.005	1.000							
LEV	0.021	0.077	0.173	0.184	0.000	0.121	0.178	0.118	0.037	1.000						
LOSS	0.004	0.000	6.960	0.288	0.032	0.004	0.000	0.010	-0.017	0.612	1.000					
OVERCON	0.059	0.913	0.017	0.114	0.014	0.005	3.670	0.004	0.070	-0.123	0.075	1.000				
ROA	0.289	0.000	0.007	0.001	0.291	0.010	0.006	0.001	-0.073	-0.018	-0.021	0.080	1.000			
SANCTION	0.015	0.016	0.168	0.090	0.001	0.311	0.056	0.000	-0.060	0.001	-0.115	0.041	-0.107	1.000		
SIZE	0.993	0.234	0.086	0.066	0.001	0.010	0.011	0.057	-0.002	-0.125	-0.025	0.913	-0.020	0.330	1.000	
TA	0.000	0.000	0.133	0.865	0.075	0.052	0.002	0.339	0.062	-0.306	0.983	-0.073	-0.055	-0.0472	-0.017	1.000

Table 5. Results of diagnostic tests for the first model

Test Name	Beta (Chi ² or F)	P-Value	Sharpe (Chi ² or F)	P-Value	Volatility (Chi ² or F)	P-Value
Poolability Test (F-Limer)	1.276	0.018**	1.258	0.025**	2.510	0.000***
Hausman Test	382.41	0.000***	34.73	0.000***	85.14	0.000***
Breusch–Pagan Test	92.672	0.000***	84.470	0.000***	177.142	0.000***

Note: The null hypotheses of the tests assume pooled data and homoskedasticity; * denotes significance at the 90% confidence level, ** at the 95% confidence level, and *** at the 99% confidence level.

As inflation rises, investors may be more hesitant to invest in risky assets, as evidenced by a negative and statistically significant correlation between inflation and the beta indicator ($p = 0.0023$). Other variables, such as firm age, board size, financial leverage, and sales growth, are largely insignificant. Overall, the results demonstrate that improvements in ESG performance are significantly associated with reductions in CMS. ESG plays an important role in mitigating systematic risk and promoting investment sustainability through lower beta values and reduced volatility. The overall model statistics indicate that the R-squared and adjusted R-squared values are moderate, suggesting that the model variables explain some of the variation in market sensitivity. At the same time, additional factors may also influence it. Moreover, the overall significance level of the model is 0.000, which is below the 5% threshold, confirming that the model is statistically significant.

Table 6. Estimation results of the first research model

Variable	beta		sharp		vol	
	Coefficient	P-Value	Coefficient	P-Value	Coefficient t	P-Value
ESG	-0.252	0.000***	-0.064	0.000***	-0.081	0.021**
OVERCON	0.026**	0.179	-1.215	0.179	-0.787	0.000***
LEV	0.233	0.203	0.874	0.203	-0.010	0.976
GRW	-0.002	0.928	0.054	0.928	0.052	0.360
SIZE	0.073	0.348	0.096	0.348	-0.272	0.009***
AGE	-0.073	0.877	0.620	0.877	0.295	0.513
ROA	0.201	0.439	1.114	0.439	-1.976	0.000***
LOSS	-0.134	0.162	0.407	0.162	-0.167	0.542
BSIZE	-0.093	0.270	-0.069	0.270	-0.078	0.395
CFO	-0.268	0.165	0.140	0.165	0.981	0.031**
TA	0.369	0.050**	0.621	0.050**	-0.433	0.191
INFLATION	-1.161	0.002***	0.677	0.002***	-0.471	0.385
SANCTION	0.343	0.000***	0.123	0.000***	1.032	0.000***
C	-0.205	0.833	-3.663	0.833	-0.457	0.578
R-squared	0.277		0.298		0.446	
Adjusted R-squared	0.171		0.196		0.365	
S.E. of regression	0.704		1.596		1.471	
Sum squared resid	581.751		2987.381		2536.131	
Log likelihood	-1344.346		-2443.811		-2333.766	
F-statistic	2.628		2.915		5.529	
Prob(F-statistic)	0.000***		0.000***		0.000***	

4.5. Additional analyses

The severity of sanctions and inflation's moderating effect on the ESG relationship and CMS were investigated in Model 2, as well as in additional analyses, the results of which are presented in Table 4.

Model 2

$$MS_{it} = a_0 + a_1ESG_{it} + a_2Inflation_{it} + a_3ESG * Inflation_{it} + a_4Sanction_{it} + a_5ESG * Sanction_{it} + a_6BSIZE_{it} + a_7age_{it} + a_8roa_{it} + a_9_{it} + a_{10}loss_{it} + a_{11}CFO_{it} + a_{12}TA_{it} + a_{13}lev_{it} + a_{14}Over.Con_{it} + a_{15}GRW_{it} + a_{16}size_{it} + a_{17}Year_{it} + a_{18}Industry_{it} + \varepsilon_{it}$$

Table 7 reports the results of the diagnostic tests for the first research model. These tests include the poolability test, the Hausman test, and the test for homogeneity of variance. To examine whether the data could be pooled, the F-Limer test was employed under the null hypothesis of pooling. Test statistics of 1.276 for beta, 2.546 for Sharpe, and 1.682 for volatility were calculated. Since the calculated F values exceed the critical values, the null hypothesis is rejected, indicating that the first model should be estimated using panel data techniques.

Subsequently, the Hausman test was applied to determine whether a fixed-effects or random-effects specification is more appropriate. The test statistics are 382.41 for beta, 84.39 for Sharpe, and 91.87 for volatility. Based on the calculated values, the null hypothesis is rejected at the 99% confidence level in favor of the random-effects model in all three cases. Therefore, the fixed-effects approach is identified as the most appropriate method for estimating this model.

To assess homoskedasticity, the regression residuals were tested with the Breusch-Pagan test. For beta (163.4), Sharpe (132.79), and volatility, the test statistics are 92.672, 163.4, and 132.79, respectively. As these values exceed the critical χ^2 values, the null hypothesis of homoskedasticity is rejected at the 99% confidence level. Accordingly, the residuals of this regression exhibit

heteroskedasticity.

Table 7. Results of diagnostic tests for the first model

Test Name	Beta (Chi ² or F)	P-Value	Sharpe (Chi ² or F)	P-Value	Volatility (Chi ² or F)	P-Value
Poolability Test (F-Limer)	1.276	0.000***	2.546	0.000***	1.682	0.000***
Hausman Test	382.41	0.000***	84.39	0.000***	91.87	0.000***
Breusch–Pagan Test	92.672	0.000***	163.40	0.000***	132.79	0.000***

Note: The null hypotheses of the tests respectively assume pooled data and homoscedasticity.

This study measured CMS using three indicators—beta (Beta), price volatility (Volatility), and the Sharpe ratio (Sharpe) and examined firms' ESG performance as the primary independent variable. To assess the effect of ESG under varying economic and political conditions, two key environmental factors were incorporated into the model, namely inflation and economic sanctions. The results indicated that ESG performance directly mitigated CMS. One of the key contributions of this study was the analysis of the moderating effects of inflation and sanctions on the relationship between ESG performance and market sensitivity. The results reported in Table 8 indicated that:

Inflation (ESG_INFLATION) weakened ESG's protective effect on CMS. The positive coefficients for Beta (0.049, $p < 0.05$), Sharpe (0.737, $p < 0.05$), and Volatility (0.286, $p < 0.05$) suggested that, under inflationary conditions, systematic risk and price volatility among firms with high ESG performance increased relative to normal conditions, thereby moderating the risk-reducing effect of ESG. This finding implied that ESG's capacity to mitigate risk was constrained in inflationary environments, underscoring the need for policymakers to account for this interaction.

Economic sanctions (ESG_SANCTION) also operated as a key moderating factor. The positive coefficients for Beta (0.091, $p < 0.05$), Sharpe (0.064, $p < 0.05$), and Volatility (0.075, $p < 0.05$) indicated that, under sanction conditions, the protective effect of ESG diminished. Even firms with high ESG performance were exposed to higher levels of systematic risk and price volatility. This result highlights the study's innovation, as it demonstrates that ESG's capacity to mitigate market risk is contingent on prevailing political and economic conditions. Among the control variables, leverage (LEV), operating cash flow (CFO), return on assets (ROA), and excessive accounting conservatism (OVERCON) exerted significant effects on CMS. Low Beta and Volatility were linked to higher liquidity and leverage, whereas ROA helped keep prices stable. In contrast, excessive conservatism increased systematic risk. Other control variables, including firm age (AGE), firm size (SIZE), and total assets (TA), did not exhibit statistically significant effects. The findings indicate that ESG is a risk-mitigation mechanism and a tool for controlling market volatility; however, its effectiveness varies with conditions, particularly during periods of inflation and economic sanctions. The results suggest that risk assessment and investment decision-making should incorporate not only firms' ESG performance but also the surrounding economic and political environment. This part of the study makes a meaningful contribution to the existing literature on ESG and capital market risk management by emphasizing the importance of contextual factors in financial decision-making.

Table 8. Estimation results of the second research model

Variable	beta		sharp		vol	
	Coefficient	P-Value	Coefficient	P-Value	Coefficient	P-Value
ESG	-0.117	0.000***	-0.248	0.043	-0.001	0.000***
INFLATION	-0.126	0.383	0.430	0.026**	-0.151	0.574
ESG INFLATION	0.049	0.038**	0.737	0.028**	0.285	0.045**
SANCTION	0.366	0.106	0.637	0.413	-0.462	0.157
ESG_SANCTION	0.090	0.018**	0.063	0.036**	0.074	0.027**
BSIZE	-0.001	0.957	0.058	0.025**	0.054	0.341
AGE	-0.257	0.272	0.176	0.588	0.994	0.029**
ROA	0.073	0.407	0.093	0.000	-0.268	0.009***
LOSS	0.189	0.525	1.067	0.276	-1.983	0.000***
CFO	-1.146	0.067*	1.585	0.056*	-1.275	0.000***
TA	-0.024	0.486	-0.021	0.592	-0.085	0.406
LEV	-0.251	0.000***	-1.211	0.000***	-0.788	0.000***
OVERCON	0.411	0.001***	0.187	0.626	1.264	0.000***
GRW	0.227	0.000***	0.856	0.246	-0.016	0.960
SIZE	-0.008	0.957	-0.339	0.137	0.291	0.403
C	-0.500	0.000***	-4.370	0.000***	-1.060	0.000***
R-squared	0.278		0.300		0.447	
Adjusted R-squared	0.171		0.196		0.366	
S.E. of regression	0.704		1.595		1.470	
Sum squared resid	581.064		2980.079		2530.869	
Log likelihood	-1343.552		-2442.167		-2332.370	
F-statistic	2.605		2.900		5.481	
Prob(F-statistic)	0.000***		0.000***		0.000***	

5. Discussion and conclusion

The Iranian capital market has faced many problems in recent years, including the harsh consequences of economic sanctions, high inflation, and economic shifts. There is a lack of empirical research in Iran examining how ESG performance affects CMS, even as ESG becomes increasingly important globally. This study found that ESG performance could be a useful tool for developing systematic risk management in emerging markets by examining the impact of ESG on Iranian enterprises' ability to weather market volatility and increase their resilience. In addition, ESG performance can be a useful tool for managing systematic risk in developing markets. This study finds the extent to which ESG mitigates CMS changes under unfavorable political and economic circumstances. Moreover, ESG has different effects on CMS depending on the severity of economic and political conditions. From this perspective, the present research extends the existing body of knowledge on ESG and capital market risk management in Iran and offers a novel analytical framework that may be utilized by investors, corporate managers, and policymakers as a practical guide for financial and investment decision-making. Therefore, this study examined the impact of ESG performance on CMS in Iran. The findings of the present study indicated that ESG performance—encompassing ESG dimensions—played a substantial role in reducing CMS. In particular, companies that performed better on ESG metrics had better Sharpe ratios, lower beta values, and less price volatility. According to these findings, companies that follow ESG principles are better able to withstand market changes and reduce their overall risk. Companies that prioritize ESG issues have been found to have a leg up in the marketplace and to achieve longer-term growth (Wu et al., 2022; Torricelli and Bertelli, 2022). Similarly, Aydoğmuş et al. (2022) emphasized the importance of investor attention to ESG-oriented firms, which aligns with the present study's evidence regarding the reduction of CMS.

One of the innovative aspects of this study was the examination of the moderating role of environmental factors, including inflation and economic sanctions, in the effect of ESG performance

on CMS. The findings indicated that inflation and sanctions may weaken the protective effect of ESG on market sensitivity. In other words, under adverse economic and political conditions, even firms with high ESG performance may be exposed to increased market volatility. These results are consistent with studies such as Zhao et al. (2018) and Bolibok (2024), which demonstrated that environmental conditions and ownership characteristics can moderate the influence of ESG on market behavior. In addition, the corporate ownership structure was found to play an important role in CMS. Overall, the findings suggest that ESG performance is a key factor in investment decision-making and in reducing firms' systematic risk. Firms with strong ESG performance benefit from greater competitive advantages, are less exposed to market volatility, and enjoy higher levels of investor confidence. Based on the empirical evidence, the research hypothesis proposing a negative and statistically significant effect of ESG performance on CMS was supported. Beyond providing further evidence of ESG's direct effect, this study also showed that contextual and environmental factors significantly moderated the relationship. Consequently, when evaluating risk and making investment decisions, investors, managers, and lawmakers should consider ESG performance and the environmental and macroeconomic conditions in which enterprises operate. Last but not least, this study adds to the body of knowledge on ESG factors in capital market risk management by showing that ESG is more than just a social or ethical metric; it's a tool for making markets less sensitive and more financially stable for businesses. Consequently, firms should regularly disclose information related to their ESG performance. Such disclosure is likely to strengthen investor confidence and reduce stock price volatility. The establishment of domestic ESG standards and adherence to international ESG frameworks may further reinforce ESG's protective effect on market sensitivity. In addition, investors and investment funds should incorporate firms' ESG performance as a core criterion for assessing investment risk and return. Allocating capital to ESG-oriented firms may help mitigate the high-risk volatility of the Iranian market and generate more stable long-term returns. At the same time, firms and investors should account for the influence of economic and political factors, such as inflation and sanctions, on the effectiveness of ESG. For example, risk-mitigation strategies, including activity diversification and prudent financial management, may help offset adverse environmental effects. In addition, a prominent limitation of this study is the use of ESG coefficients from Fakhari et al. (2018), which may not fully reflect recent changes across all industries.

References

1. Akgun, O. T., Mudge, T. J. and Townsend, B. (2021). How company size bias in ESG scores impacts the small cap investor. *The Journal of Impact and ESG Investing*, 1(4), pp. 31-44.
2. Aldieri, L., Amendola, A. and Candila, V. (2023). The impact of ESG scores on risk market performance. *Sustainability*, 15(9), p. 7183. <https://doi.org/10.3390/su15097183>
3. Aydoğmuş, M., Gülay, G. and Ergun, K. (2022). Impact of ESG performance on firm value and profitability. *Borsa Istanbul Review*, 22, pp. S119-S127. <https://doi.org/10.1016/j.bir.2022.11.006>
4. Bakhtiari, M. (2024). Investigating the impact of environmental, social, and governance (ESG) reporting on investor risk perception. *Applied Research in Financial Reporting*, 12(2), pp. 73-102 (In Persian).
5. Bermejo Climent, R., Garrigues, I. F. F., Paraskevopoulos, I. and Santos, A. (2021). ESG disclosure and portfolio performance. *Risks*, 9(10), p. 172. <https://doi.org/10.3390/risks9100172>
6. Bissoondoyal-Bheenick, E., Brooks, R. and Do, H. X. (2023). ESG and firm performance: The

- role of size and media channels. *Economic Modelling*, 121(1), A. 106203. <https://doi.org/10.1016/j.econmod.2023.106203>
7. Bolibok, P. M. (2024). Does firm size matter for ESG risk? Cross-sectional evidence from the banking industry. *Sustainability*, 16(2), p. 679. <https://doi.org/10.3390/su16020679>
 8. Boshnak, H. A. (2025). Ownership structure and firm performance: evidence from Saudi Arabia. *Journal of Financial Reporting and Accounting*, 23(4), pp. 1656-1676. <https://doi.org/10.1108/JFRA-11-2022-0422>
 9. Devalle, A., Fiandrino, S. and Cantino, V. (2017). The linkage between ESG performance and credit ratings: A firm-level perspective analysis. *International Journal of Business and Management*, 12(9), pp. 53-65. <https://dx.doi.org/10.5539/ijbm.v12n9p53>
 10. Fakhari H., Malekian E. and Jafaei Rahni M. (2018). Explanation and ranking of environmental, social and corporate governance reporting components and indicators using the analytic hierarchy process in companies listed on the stock exchange, *Value and Behavioral Accounting*, 1(1), pp. 153-187.
 11. Hu, W., Sun, J., Lin, Y. E. and Hu, J. (2023). ESG and investment efficiency: the role of marketing capability. *Sustainability*, 15(24), A. 16676. <https://doi.org/10.3390/su152416676>
 12. Landi, G. and Sciarelli, M. (2019). Towards a more ethical market: the impact of ESG rating on corporate financial performance. *Social Responsibility Journal*, 15(1), pp. 11-27. <https://doi.org/10.1108/SRJ-11-2017-0254>
 13. Nasdaq. (2019). ESG reporting guide 2.0, A support resource for companies, *Nasdaq*, New York City, New York. pp. 1-38.
 14. Pu, G. (2023). A non-linear assessment of ESG and firm performance relationship: evidence from China. *Economic research-Ekonomska istraživanja*, 36(1), p. 2113336. <https://doi.org/10.1080/1331677X.2022.2113336>
 15. Rahimzadeh, F., Jamali, J. and Abed, A. (2022). The effect of company ownership structure on the stock price crash risk in Iran: a panel co-integration approach with cross-sectional dependence. *Iranian Journal of Accounting, Auditing and Finance*, 6(4), pp. 1-14. https://ijaaf.um.ac.ir/article_42122.html
 16. Rahman, M. A. and Hossain, M. M. (2023). Impact of ownership structure and board size on firm performance: A study on listed companies of Dhaka Stock Exchange. Noakhali Science and Technology University, Bangladesh. Available at SSRN 4527887. <https://dx.doi.org/10.2139/ssrn.4527887>
 17. Sabouri, N. and Allahyari, A. (2023). The relationship between environmental, social, and governance (ESG) reputational risk and financial performance considering regulatory dynamics. *Journal of New Research Approaches in Management and Accounting*, 7(88), pp. 1595-1611 (In Persian).
 18. Sassen, R., Hinze, A. K. and Hardeck, I. (2016). Impact of ESG factors on firm risk in Europe. *Journal of Business Economics*, 86(8), pp. 867-904. <https://doi.org/10.1007/s11573-016-0819-3>
 19. Schrand, C. M. and Zechman, S. L. (2012). Executive overconfidence and the slippery slope to financial misreporting. *Journal of Accounting and Economics*, 53(1-2), pp. 311-329. <https://doi.org/10.1016/j.jacceco.2011.09.001>
 20. Shan, X., Song, Y. and Song, P. (2025). How ESG performance impacts corporate financial performance: a DuPont analysis approach. *International Journal of Climate Change Strategies and Management*, 17(2), pp. 1-24. <https://doi.org/10.1108/IJCCSM-07-2024-0125>
 21. Torricelli, C. and Bertelli, B. (2022). ESG screening strategies and portfolio performance: how do they fare in periods of financial distress?. Work. Paper, Giugno, Università di Modena e

- Reggio Emilia. p. 87. https://dx.doi.org/10.25431/11380_1291995
22. Van Duuren, E., Plantinga, A. and Scholtens, B. (2016). ESG integration and the investment management process: Fundamental investing reinvented. *Journal of Business Ethics*, 138(3), pp. 525-533. <https://doi.org/10.1007/s10551-015-2610-8>
23. Wu, S., Li, X., Du, X. and Li, Z. (2022). The impact of ESG performance on firm value: The moderating role of ownership structure. *Sustainability*, 14(21), A. 14507. <https://doi.org/10.3390/su142114507>
24. Yang, X., Bao, J. and Zhang, K. (2025). Environmental, social and governance (ESG) performance and abnormal positive tone. *Sustainability Accounting, Management and Policy Journal*, 16(2), pp. 351-388. <https://doi.org/10.1108/SAMPJ-01-2024-0045>
25. Zhao, C., Guo, Y., Yuan, J., Wu, M., Li, D., Zhou, Y. and Kang, J. (2018). ESG and corporate financial performance: Empirical evidence from China's listed power generation companies. *Sustainability*, 10(8), p. 2607. <https://doi.org/10.3390/su10082607>
26. Zhou, G., Liu, L. and Luo, S. (2022). Sustainable development, ESG performance and company market value: Mediating effect of financial performance. *Business Strategy and the Environment*, 31(7), pp. 3371-3387. <https://doi.org/10.1002/bse.3089>



RESEARCH ARTICLE

The Effectiveness of Kahoot! Game in Teaching Accounting Principles Course on Students' Quantitative and Qualitative Academic Performance

Mohammad Mahdi Bozorg

Finance-Financial Engineering, Yazd University, Yazd, Iran

Zohreh Arefmanesh*

Department of Accounting and Finance, Faculty of Economics, Management, and Accounting, Yazd University, Yazd, Iran

Ahmad Zandvanian

Department of Education, Yazd University, Yazd, Iran

How to cite this article:

Bozorg, M. M. , Arefmaesh, Z. and Zandvanian, A. (2026). The Effectiveness of Kahoot! Game in Teaching Accounting Principles Course on Students' Quantitative and Qualitative Academic Performance. Iranian Journal of Accounting, Auditing and Finance, 10(3), 113-137. doi: 10.22067/ijaaf.2026.48263.1634
https://ijaaf.um.ac.ir/article_48263.html

ARTICLE INFO

Article History

Received: 2025-12-10

Accepted: 2026-03-25

Published online: 2026-07-12

Keywords:

Accounting Education,
Educational Technology,
Game-based Learning,
Gamification, Kahoot

Abstract

This study investigates the use of the educational technology tool Kahoot! in Accounting Principles II within a quasi-experimental design, addressing limited empirical evidence regarding students' perceptual and performance-related responses to such tools in accounting education. Although various digital platforms have been explored in higher education, the documented effects of these platforms on students' perceptions and performance in accounting courses remain relatively limited. Here, we extended this evidence by dissociating perceptual and performance-related measures. This study involves 63 undergraduate accounting and financial management students from Yazd University, who were divided into two groups: an experimental group that used Kahoot! for in-class exercises and assessments, and a control group that completed similar activities on paper. Data were analyzed using Excel and SPSS 27, with ANCOVA, MANCOVA, and a one-sample t-test used. Students in the experimental groups shared positive beliefs about motivation, enjoyment, collaboration, and classroom engagement. Yet when academic performance between the two waves was statistically controlled for baseline achievement, no statistically significant differences were found; initial within-group nominal differences in some classroom tests failed to survive Bonferroni adjustments. Because of the quasi-experimental design with nonequivalent groups, modest sample size, and low statistical power, these findings reflect observed associations rather than causal relationships and must be interpreted with caution. Overall, while Kahoot! While practicing at home may allow for a much more dynamic classroom experience, we cannot be certain whether it impacted their test performance. This study offers a timeline-based framework to inform future research on the pedagogical use of digital instructional tools in accounting education.

<https://doi.org/10.22067/ijaaf.2026.48263.1634>



NUMBER OF REFERENCES
64



NUMBER OF FIGURES
1



NUMBER OF TABLES
9

Homepage: <https://ijaaf.um.ac.ir>
E-ISSN: 2717-4131
P-ISSN: 2588-6142

*Corresponding Author: Zohreh Arefmanesh
Email: arefmanesh@yazd.ac.ir
Tel: 09133577956
ORCID: 0000-0001-7328-2035

1. Introduction

As the demands of the modern academic environment shift, students want more engaging group interactions than one-way lecturing. For instance, in fields like accounting, which are often taught from a very rigorous, theoretical standpoint, this change is particularly noticeable, as many students do not benefit from static learning when confronting the abstract nature of foundational content. With the digital generation occupying most university classroom seats, higher education is grappling with how to reconfigure pedagogies, as students today have different learning styles.

Previous research has highlighted that the use of technology in higher-education accounting teaching contexts is associated with greater student engagement and motivation (Go et al., 2024) and that the use of game elements in teaching may be related to a better understanding of accounting concepts (Ramesh et al., 2025). Recently, game-based learning (GBL) and gamification have emerged as promising approaches for the modern instructional experience, combining interactive elements with enjoyable activities that mirror students' digital habits and are readily available for educational purposes (Neerupa et al., 2024). Theoretically, gamification in the educational context is usually justified by Goal-Setting Theory and Expectancy Theory, which emphasize that clearly defined goals are more effective when accompanied by prompt feedback and when students perceive a strong association between effort, performance, and valued outcomes. Gamification in accounting education features student motivation and engagement, autonomy, competence, and relatedness needs. Metricated positive impacts on exam scores and problem-solving abilities. Mechanisms for accessing these effects. Resource limitations & gap inclination behavioral Gap issues of student diversity Nubial Outcome Anwar and Mulyadi (2025) Platforms such as Kahoot! are known to increase student participation within the classroom by allowing immediate feedback in a low-pressure, game-like environment. However, despite some positive impact of gamification on student performance, the relationship between gamification and academic performance is so complex that findings from previous studies are highly mixed (Bhavani et al., 2020), and more research in these areas, especially in unexplored educational contexts such as Iran, is necessary.

Over the years, several studies have investigated the use of Kahoot! Most of the research has centered on student perceptions, engagement, and motivation rather than on externally validated learning outcomes (e.g., in accounting-specific courses) and on measured competencies appropriate to disciplines (which warrant further statistical analysis). Nevertheless, we have identified only a handful of studies that directly measure the association of Kahoot! They are founded on formative assessment and key quantitative performance indicators in accounting lessons. The current study seeks to contribute to this burgeoning body of literature by investigating this relationship in the context of Principles of Accounting II (a second-level required accounting course) at a single Iranian university. Furthermore, the study contributes to the understanding of real academic performance rather than self-reported perceptions by using a quasi-experimental design with a nonequivalent control group. In addition, based on Goal-Setting Theory Kahoot!, implements structured quizzes and challenges by points With respect to the first, Clear Goals performance feedback can clarify learning goals and enhance students' engagement with these goals (Durso et al., 2020) while Expectancy Theory implies that transparent scoring and leaderboards may bolster students' motivation by reinforcing beliefs that increased effort leads to improved performance, leading to desirable academic outcomes (Chan et al., 2021). However, the cultural perception of using smartphones in class and in educational games is new to students, so their academic behavior may not be well understood. To fill these gaps, this study analyzed the correlation between integrating Kahoot! into Accounting Principles II and students' academic performance, in terms of quantitative outcomes (grades) as well as qualitative ones (perceptions and motivation). This study uses a quasi-experimental method. The research population consists of 63 undergraduate students who have studied accounting and are

studying financial management at Yazd University. In the Kahoot! condition, please note that the experimental group completed all in-class exercises and quizzes, while the control group continued with traditional paper-based assessments.

2. Literature review

Even if there are studies ([Pitic and Irimiaș, 2023](#)) that show the positive impact of educational games on learning in higher education, these studies mainly focus on pre-university students. Undergraduate studies have identified them as elements that encourage active and self-regulated learning, as well as collaboration ([Ayllón et al., 2019](#); [Wanner and Palmer, 2018](#); [Bovill, 2020](#)). The interactivity of these online educational games can help improve knowledge retention and conceptual understanding compared with more traditional paper-based activities. Furthermore, learning from mistakes in gamified contexts fosters deeper cognitive reinforcement and continuous development ([Jääskä and Aaltonen, 2022](#)). Given the typically low participation rates in accounting classes ([Nkhoma et al., 2018](#)), the use of such interactive technologies may enhance both engagement and academic performance. These outcomes operate across two levels: individual-level cognitive motivational processes (e.g., effort, perceived value) and classroom-level environmental factors (e.g., feedback immediacy, competitive climate), which collectively shape performance.

2.1. Gamification

Gamification, defined as the use of game elements in non-game contexts ([Basten, 2017](#)), provides a novel way to increase intrinsic motivation, satisfaction, and learning performance ([Montola et al., 2009](#); [Mekler et al., 2013](#)). It generates motivation through competition, accomplishment, and feedback. On the other hand, competition can and does lead to positive outcomes; however, excessive emphasis on competition can lead to negative outcomes, such as learner frustration ([Zhang, 2008](#)). Gamification has been shown to support theoretical retention and long-term engagement among technology-driven learning strategies ([Aguilar-Castillo et al., 2021](#)). To date, gamification is theoretically grounded in Goal-Setting Theory ([Locke and Latham, 2002](#)), which argues that specific, achievable goals increase motivation and task performance. In classroom contexts, points and badges and even leaderboards are incorporated into games as part of gamified tools like Kahoot! behaving as proximal goals that can motivate persistent effort and engagement. Goal-related feedback is linked to self-evaluation and achievement striving, thereby reinforcing learners' commitment to academic achievement. The detailed mechanisms described are most relevant to Generation Z students who habitually use mobile devices due to their digital fluency and their preference for mobile-supplemented or blended learning formats ([Vizcaya-Moreno and Pérez-Cañaveras, 2020](#)), thus organizing positively within a critical structure with goal-oriented outputs.

In addition, Expectancy Theory ([Vroom, 1964](#)) provides a complementary psychological explanation for gamification effects. According to this model, motivation is driven by three perceptions: expectancy (the belief that effort improves performance), instrumentality (the belief that performance leads to rewards), and valence (the value attached to those rewards). Gamified quizzes provide instant feedback, thereby generating a sense of expectancy; visible scoreboards promote instrumentality and a fun interface elevates the perception of valence. In summary, the mechanisms help explain how gamification can increase or clarify goals, sustain engagement, and ultimately improve the academic performance of digitally born cohorts, such as Generation Z accounting students.

2.2. Gamification and psychological levels in accounting education

The variables in this study can be understood through three distinct psychological levels, which help bridge the gap between gamified interventions and academic success in accounting:

1. *Cognitive Level*: Technology-enhanced learning tools, such as real-time online assessments, offer the instant feedback needed to improve knowledge retention and metacognitive awareness (Chappuis et al., 2020; Attali and Van Der Kleij, 2017). Such tools are associated with deeper processing and conceptual understanding in the domain of accounting, a complex field (Krasodomska and Godawska, 2021).
2. *Motivational Level*: Competitive architecture of gamification, specifically point systems and leaderboards, is utilized to invoke goal commitment and perceived competence (Göksün and Gürsoy, 2019; Sailer and Homner, 2020). Recent empirical research in accounting education also finds that these motivational triggers are strongly related to students' academic performance (López-Hernández et al., 2023).
3. *Behavioral Level*: Interactive participation during Kahoot! Sessions increase student attention and classroom interaction, effectively addressing the historically low participation levels observed in traditional accounting instruction (Nkhoma et al., 2018). As highlighted by Lew and Saville (2021), such game-based learning experiences are essential for transforming passive listeners into active participants, thereby fostering a more robust behavioral engagement with technical accounting content.

The theoretical integration elucidates how gamified assessments affect learning across multiple psychological levels by mapping these variables to the cognitive, motivational, and behavioral domains. This multifaceted view provides an important response to a previous criticism of the psychology literature, as many works often ignore how the interplay of psychological domains among students affects objective academic performance.

Use of Kahoot! as a Gamified Learning Tool: Kahoot! is a popular web-based platform used for formative assessments through interactive quizzes (Wang and Tahir, 2020). As of 2024, Kahoot! had hosted more than 10 billion participants, offered over 50 million publicly available games, and attracted over 1 billion unique players annually (Retrieved from Kahoot! website, <https://kahoot.com>, July 1, 2025). These figures indicate not only widespread adoption but also the platform's perceived usefulness across diverse educational levels and disciplines. In higher education, Kahoot! is frequently used to review course content, check conceptual understanding, and promote participation during lectures. It practically operationalizes the principles of both Goal-Setting and Expectancy Theories by providing structured challenges, transparent scoring, and immediate performance feedback. The tool's engaging and competitive format has been associated with enhanced motivation and performance across diverse disciplines, although mixed findings exist regarding learning outcomes (Lee et al., 2019; Wang et al., 2016; Ranieri et al., 2021). In accounting education, Kahoot! Has been linked with measurable increases in learning activity and post-test scores (Wardana and Sagoro, 2019). Similarly, international studies have emphasized its motivational and formative benefits (Gómez and Menroy, 2018; Ortiz-Martínez et al., 2022). These findings are particularly relevant to accounting, given students' greater need for feedback-driven practice to master abstract numerical concepts.

Synthesis of Kahoot! Impacts: Learning, Dynamics, and Affective Responses: The impact of Kahoot! on well-being, and academic achievement is associated with substantial motivational benefits, but its impact on objective school performance depends on the quality of implementation. While broad consensus links Kahoot! to increased concentration and engagement (Lee et al., 2019; Ortiz-Martínez

et al., 2022), findings regarding final achievement are mixed, with some evidence suggesting that effectiveness may be limited in complex or hands-on subjects (Wang et al., 2016; Ranieri et al., 2021). In the context of accounting education—which demands a blend of conceptual and procedural rigor—studies such as Wardana and Sagoro (2019) have demonstrated substantial gains in both learning activity and test scores, reinforcing that when Kahoot! is strategically aligned with curriculum objectives, it functions as a powerful tool for formative assessment and knowledge retention (Bakar, 2024; Omar et al., 2024).

Beyond performance outcomes, the effectiveness of Kahoot! lies in its ability to reshape classroom dynamics and manage affective learner responses. By transitioning instruction toward a student-centered model, the platform enhances interaction and reduces classroom-related stress, particularly for reserved learners who may otherwise be reluctant to participate (Bawa, 2019; Senali et al., 2022). Although the competitive nature of Kahoot! has raised concerns regarding performance anxiety and rushed responses (Wang and Tahir, 2020; Göksün and Gürsoy, 2019), recent literature highlights that this competition, when well-managed, fosters sustained involvement and positive attitudes (Turan and Meral, 2018; Almuntsr et al., 2024; Fortuna et al., 2023). Importantly, both students and teachers agree that the platform is useful because it reduces boredom and workload, but recent calls for research emphasize that more work is needed to understand how gamified interventions can address the multitude of diverse needs (Elkhamisy and Wassef, 2021; Ionescu-Feleagă et al., 2025).

Use of Educational Games in Accounting Education in Iran: Despite this global contribution to the review of gamification, a new review of gamification in accounting education has shown how game-based approaches can enhance student motivation, engagement, and learning across different accounting courses (Queiro Ameijeiras et al., 2025). Evidence, however, is scarce and descriptive in the Iranian higher-education context, but Haji Moradkhani et al. (2017) found in the first study, conducted at the University of Tehran, it was found that teaching with board games such as Monopoly significantly improved student engagement and learning outcomes compared with traditional didactic methods (2018). Alternatively, Khatibi and Badeleh (2021) found that in other Iranian educational contexts, Kahoot! Drawing from a computational perspective, research found that game-based activities increased learner engagement, language skills, and overall satisfaction with the learning process; this work typically describes motivational and perceptual outcomes without concern for the magnitude of the impact on course grades. The most relevant information was found in Devanipour (2022), which studied Kahoot! Formative assessment in a Cost Accounting course Improvements in questionnaire data, such as enjoyment and perception of engagement compared to the traditional paper-based assessments were observed; however no differences in mean final examination scores between Kahoot! was not significant between the gamified instruction and control groups, indicating that the relationship of gamified instruction with academic performance seems to be context-, and more importantly perhaps, learning experience-specific (regarding course design, content, and technological conditions). Overall, these results suggest that classroom-based educational games may facilitate more dynamic and participatory lessons in accounting classrooms; however, there is a lack of empirical studies on digital platforms such as Kahoot! in Iranian university settings.

Summary and Critical Reflections: About Kahoot! In summary, the literature indicates that Kahoot! is commonly associated with improved student motivation, engagement, and classroom dynamics at the college level. Though the platform has proven particularly effective at reinforcing theory and review-based knowledge, its impact on complex, applied, or skills-based learning is still hotly debated. Moreover, the competitive framework innate to Kahoot! This does function as a double-edged sword: the mechanism it uses to encourage participation works, whilst also rearing

hyper-competitive performance anxiety in some learners, meaning the pedagogical promise is highly dependent on the implementation strategies chosen.

However, many findings are hindered by limited longitudinal and cross-cultural research, which limits their generalizability across diverse educational contexts. On a methodological level, there is a lack of studies that account for differences in learner characteristics, such as prior academic performance and learning styles. These gaps require enhanced empirical examination, particularly quasi-experimental and mixed-methods designs to move beyond descriptions and better evaluate the educational potential of game-based platforms in teaching and learning. This study overcomes some of these limitations by moving the focus to a more objective analysis of gamified formative assessment. Kahoot! This study provides important, context-dependent empirical evidence on how accounting knowledge affects both objective academic performance and subjective student perceptions in the undergraduate accounting context in Iran. This paper contributes to the wider gamification literature by linking Goal-Setting and Expectancy theories to explain how digital interventions might be aligned with learning outcomes for accounting students.

3. Methodology

3.1. Population, sample, and sampling method

This study employed a quasi-experimental design with a nonequivalent control group. The statistical population comprised undergraduate students enrolled in the Principles of Accounting II course at Yazd University during the 2022–2023 academic year. The sample consisted of two intact classes: 36 Accounting students in the experimental group and 30 Financial Management students in the control group. Although these groups belonged to different majors and therefore may have differed in prior academic background, both programs shared a common foundational curriculum in Principles of Accounting. This shared prerequisite provided a meaningful basis for comparison in the present study. To account for possible baseline differences in prior knowledge, the students' scores in Principles of Accounting I were entered as a covariate in the statistical analysis (ANCOVA). This approach allowed the post-intervention outcomes to be adjusted for initial differences, thereby strengthening the internal validity of the comparison. Nevertheless, the distinct nature of the two majors should still be considered as a contextual limitation in interpreting the findings.

The same researcher-instructor taught both groups to ensure consistency in instructional delivery. The experimental group was taught via Kahoot! The experimental group used the interactive learning method, while the control group used the paper-based assessment method. Protocols were standardized with the following procedures to minimize expectancy effects and differential treatment:

1. *Strict Protocol Adherence:* Both groups used the same textbook, followed the same instructional schedule, and attended the same number of sessions.
2. *Identical Assessment Content:* The study employed Functional Phased Tests to monitor progress throughout the semester. In addition, identical Midterm and Final examinations were administered to both groups to ensure content equivalence and comparability.
3. *Anonymized Grading:* Written answer sheets for the Midterm and Final examinations were graded anonymously. Student names were removed before evaluation to minimize any conscious or unconscious bias.
4. *Automated Data Collection:* Kahoot! Scores were automatically recorded and exported directly from the platform, with no manual intervention in data extraction. This procedure enhanced objectivity and transparency in the measurement process.

Measures and Data Collection: The main outcome measurements comprised the Functional Phased Tests score, scores on the Midterm and Final examinations, and performance in the Kahoot!

Quiz scores. The measure of learning was developed to closely align with the course objectives and the content delivered in class over a semester.

To support the validity and reliability of the measurement process, the following steps were taken:

- *Content Validity:* The examinations were prepared by the instructor and reviewed by three faculty experts in accounting education to ensure alignment with the course syllabus and learning outcomes.
- *Reliability:* The internal consistency of the summative assessments was examined using Cronbach's alpha. The coefficient indicated acceptable reliability for the instruments used in the study.
- *Baseline Covariate:* The score from Principles of Accounting I was used as a baseline covariate rather than a conventional pre-test. This decision was made to statistically adjust for prior knowledge without introducing testing effects that may arise from administering an identical pre-test.

Limitations: This study has several limitations that should be acknowledged. First, the use of nonequivalent groups means that random assignment was not possible, which may introduce selection bias. Although ANCOVA was employed to statistically control for baseline differences, this limitation cannot be eliminated. Second, the participants were students in this case, from different majors, but, arguably, their common introductory accounting course and a covariate made this less worrisome. Finally, the researcher served as the instructor, but procedural safeguards such as standardizing instruction, anonymizing grading, and automating data collection were implemented to mitigate bias. Finally, the sample size was constrained by the availability of intact classes, which is a common limitation in educational quasi-experiments.

3.2. Research hypotheses

H1: Students hold a positive perception regarding their experience with the Kahoot! Platform for teaching Principles of Accounting II.

Rationale: This hypothesis stems from the Expectancy Theory, where the gamified features of Kahoot! (e.g., immediate performance feedback and competitive leaderboards) act as positive valence that elevates students' expectations of success and engagement, directly influencing their overall perception of learning.

H2: There is a statistically significant difference in academic performance in Principles of Accounting II between students who used Kahoot! And those who received traditional instruction.

Rationale: Grounded in Goal-Setting Theory, this hypothesis posits that Kahoot! Provides a structured environment that facilitates clear, proximal learning goals. We define "academic performance" as the mastery of accounting principles; we hypothesize that the continuous, goal-oriented reinforcement provided by Kahoot! Leads to higher achievement compared to traditional, less interactive instruction.

H3: There is a statistically significant difference in classroom test performance between students who received Kahoot!-based formative assessment and those who received traditional paper-based formative assessment.

Rationale: This is linked to the cognitive mechanisms of Goal-Setting Theory. Kahoot! replaces traditional paper-based formative assessments. The intervention creates a closed-loop system. We conjecture that this immediate cognitive feedback loop allows students to target their studies more effectively, thereby producing greater academic gains on classroom examinations than students who use distant, paper-and-pencil evaluations.

3.3. Data collection tools

3.3.1. Researcher-Made Questionnaire

A questionnaire was developed to gather student perspectives on learning effectiveness. It was based on validated instruments from [Lin et al. \(2018\)](#), [Silva et al. \(2019, 2021\)](#), [Marsa et al. \(2021\)](#), and [Haji Moradkhani et al. \(2018\)](#). Content validity was established through expert review by educational researchers. Reliability was assessed using Cronbach's alpha and McDonald's omega. For the two-item subscale, *Value of using Kahoot in the classroom*, Cronbach's alpha was $\alpha = 0.65$, and McDonald's omega was $\omega = 0.65$. Although slightly below the conventional threshold of 0.70, these values are acceptable for very short scales, as reliability coefficients are highly sensitive to item count and tend to underestimate internal consistency in two-item measures. More importantly, both items loaded highly on a single factor (0.861), indicating considerable shared variance and unidimensionality of the construct. To provide additional evidence regarding construct validity, an Exploratory Factor Analysis (EFA) was carried out on the entire 25-item questionnaire. The Kaiser–Meyer–Olkin (KMO) measure of sampling adequacy was 0.78 (>0.60), indicating that the factor analysis results are reliable [21]. Bartlett's test of sphericity was statistically significant $\chi^2(300) = 742.36$, $p < 0.001$ or examination of scree plots, leading to a structure theoretically consistent with proposed dimensions. Given the relatively small sample size, the EFA findings were interpreted with caution and used only for preliminary construct verification, not for full structural validation. This approach is supported by methodological literature indicating that when EFA is used for exploratory verification—especially with instruments adapted from previously validated questionnaires—strong factor loadings provide meaningful evidence of construct representation even in smaller samples ([Fabrigar et al., 1999](#)). For the entire 25-item instrument, Cronbach's alpha was 0.91, indicating excellent internal consistency (see Table 1 for details on dimensions and item counts).

Table 1. Dimensions of researcher-made questionnaire

Dimensions	Number of items	Item numbers	Cronbach's alpha
The value of using Kahoot in the classroom	2	1-2	0.65*
The effectiveness of Kahoot in understanding the course materials	6	3-4-5-6-7-8	0.83
Benefits of using Kahoot inside the classroom	5	9-10-11-12-13	0.81
Elements of Kahoot! itself	12	14-15-16-17-18-19-20-21-22-23-24-25	0.82

* **Note:** For the two-item subscale, Cronbach's alpha is 0.65. However, this reliability estimate is corroborated by McDonald's omega ($\omega = 0.65$) and by strong construct validity, as confirmed via EFA, with both items loading highly on a single factor (0.861). These converging results support the adequacy of this dimension.

A 5-point Likert scale was used, ranging from Strongly Agree (5) to Strongly Disagree (1). The questionnaire was conducted during the final Kahoot! session with 36 accounting students on an anonymous basis. The session in which they were assured that responses would not affect their course grades.

Design logic and structural alignment with Hypothesis 1: The survey was designed to operationalize Hypothesis 1, which posits student perceptions and involvement during the Kahoot! experience. Compared with traditional instruction. To holistically measure this construct, the instrument was divided into four correlated dimensions representing various phases of the learning experience:

1. *Value of Kahoot (Items 1–2)*: Assesses the perceived usefulness and relevance of Kahoot in the classroom, reflecting students' evaluative appraisal of the tool.
2. *Effectiveness in Understanding (Items 3–8)*: Measures students' perceptions of how Kahoot

supports understanding of course materials.

3. *Classroom Benefits (Items 9–13)*: Evaluates motivational and engagement-related benefits associated with Kahoot use.
4. *Elements of Kahoot (Items 14–25)*: Examines specific game mechanics (e.g., points, competition, and feedback) that may contribute to instructional effectiveness.

3.3.2. Functional phased tests

These were formative classroom assessments focused on chapters and short sections, enhancing learning and overcoming hurdles. It should be noted that, in this work, Kahoot was used as a formative assessment tool rather than as direct instruction. The intervention involved using Kahoot! Scheduled competitive review sessions were held at the end of every topic, which gave students a chance to re-evaluate their understanding and gain instant feedback before major tests. This is different from what we would call "game-based learning", where the game itself serves as the primary medium of instruction. Instead, this research assesses the relationship between game-based formative assessment and student performance on classroom tests, compared with traditional paper-based formative assessments. They included stage-by-stage tests throughout the semester to spur participation and discussed regular practice by framing segments.

Control group: Used traditional paper-based evaluations.

Experimental group: Answered similar questions via Kahoot!

Exam topics: Inventory, cash, payroll, partnership I, and stock company.

3.3.3. Midterm and final examinations

The two groups were given the same midterm and final examinations on the same dates under similar conditions. The last test was on partnerships and corporations, with students in both classes answering the same questions at the same time.

3.3.4. Additional research variables

Additional student data were collected, including gender, birthdate, high school major, GPA, entrance exam rank, guest student status, major change, pre-test score for Accounting Principles I, and prior history of failing Accounting Principles I.

3.3.5. Research implementation process

Throughout the semester, accounting students prepared for Kahoot! Quizzes at the end of each topic, using their mobile phones to answer questions on the Kahoot! Online platform. The system reports were useful for obtaining data for analysis. The financial management students also followed the same scheduled reviews and tests as before. Still, instead of computer-based methods, they used paper-based methods to answer all questions, with both groups in each test taking the same content.

1. Intervention program – Lesson Plan of Accounting Principles II

Table 2 presents the lesson plan for Accounting Principles II, highlighting the use of Kahoot! Game. It outlines the required sessions for teaching each topic and the activities for each session.

Table 2. Lesson plan for Accounting Principles II in the experimental and control groups

Session	Activities	Experimental Group	Control Group
1	Inventory Chapter	*	*
2	Fixing the mistakes and Solving Exercise - Cash Chapter	*	*
	The first functional step test – the topic of inventory (15 test questions)	Kahoot	Paper Assessment
3	Cash Accounting (Preparation of Bank Reconciliation Statement)	*	*
4	Solving Cash Accounting Exercise - Payroll Chapter	*	*
	Second Functional Stage Test – Cash Topic (15 Test Questions)	Kahoot	Paper Assessment
5	Solving Exercises for Payroll Accounting: How to Record Payroll Expense	*	*
6	Partnership Chapter: Types of Accounts and How to Register Them, and How to Divide Profits and Losses	*	*
	Third Functional Stage Test – Salary Topic (15 Test Questions)	Kahoot	Paper Assessment
	Mid-Term Exam of Inventory, Cash, and Payroll (Descriptive)	Paper Assessment	Paper Assessment
7	Preparing Financial and Accounting Statements for New Partner Entry	*	*
8	Exit of the Partner in the Partnership	*	*
	Fourth Functional Stage Test: Partnership Topic 1 (Dividend Distribution, Entry of a New Partner - 10 Test Questions)	Kahoot	Paper Assessment
9	Solving Exercise - Partnership Accounting: Liquidation	*	*
10	Stock Company Chapter: Stock Release	*	*
	Fifth Functional Stage Test: Partnership Topic 2 (Partner Exit and Instant Dissolution - 10 Test Questions)	Kahoot	Paper Assessment
11	Stock Company Formation Process	*	*
12	Solving Exercises – Profit distribution, year-end operations, and preparation of financial statements in stock companies	*	*
	Sixth Functional Stage Test – Stock Company Topic (15 Test Questions)	Kahoot	Paper Assessment
	Final exam on the topics of partnership and stock company (descriptive)	Paper Assessment	Paper Assessment

At the end of each topic, accounting students were given one week to study and summarize the material before competing in a live Kahoot! Session in class. Meanwhile, financial management students answered similar paper-based assessments under the same study conditions. Due to holiday breaks and scheduling constraints, the sixth functional stage exam was omitted, though stock company topics were included in the final exam without a separate study week.

3.3.6. Collecting educational questions

Multiple-choice questions and the standard Kahoot! Formats were carefully selected for educational value and alignment with course content. Questions were sourced from the Principles of Accounting II textbook and national exams, avoiding newly designed questions. The instructor reviewed and refined them before entry into Kahoot!.

3.3.7. Implementation of Kahoot!

Each Kahoot! Was custom-built for its topic on Kahoot!'s website. Two usage modes exist: live in class (chosen for this study) or independent practice. To ensure a smooth game session:

- Questions were pre-entered and previewed on the platform.
- Students were informed in advance to review topics and check device battery levels.
- The session required a video projector, a wide screen, and stable internet access for all participants.
- The classic mode was selected for live gameplay and evaluation.

To participate, students accessed Kahoot! Log in to your mobile or computer, enter the game code, and check the real-time leaderboard. No account creation was necessary. The teacher started the live in-class competition, and students answered a question within a limited timeframe. We provided the correct answers after each round, which enabled immediate feedback and discussion.

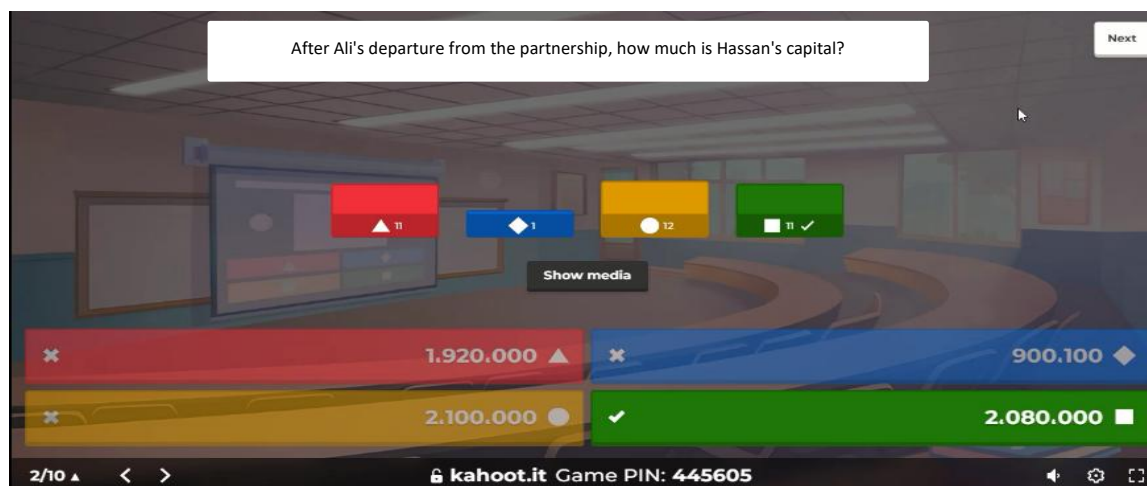


Figure1. The game screen, when displaying the answer to one of the questions

- Scores were ranked based on accuracy and speed, fostering a competitive learning environment.
- The top five players were displayed after the game, with individual results viewable on devices.
- The instructor reviewed a real-time summary, identifying challenging questions for further analysis.
- Players could provide session feedback via Kahoot!'s built-in tool.

3.3.8. Data analysis method

Data were processed using Excel and SPSS 27.

Principles of Accounting I scores were used as a pre-test measure.

Final scores for Principles of Accounting II served as the post-test evaluation.

Statistical methods applied:

- Analysis of covariance (ANCOVA)
- Multivariate analysis of covariance (MANCOVA)
- One-sample t-test

4. Results

The results are presented in relation to the study hypotheses. First, the descriptive findings are summarized, followed by the inferential analyses used to test each hypothesis. As shown in Table 3, the study included 63 undergraduate students: 36 (57.1%) in the experimental group (accounting) and 27 (42.9%) in the control group (financial management). Because intact classes were used, the groups differed in both size and academic major, which may introduce pre-existing differences that statistical controls cannot fully eliminate. The final sample consisted of 63 students, of whom 45 (71.4%) were female, and 18 (28.6%) were male. Initially, the control group comprised 30 financial management students; however, the final count was reduced to 27 after one student withdrew before the midterm exam, another before the final exam, and a third dropped the Principles of Accounting II course after

missing the final exam. Participants ranged in age from 18 to 24 years and, having been born between the late 1990s and mid-2010s, are classified as members of Generation Z.

Table 3. Gender and age range of participants by group

Participant characteristics		Experimental Group (Accounting)	Control Group (Financial Management)
Gender	Female	29	16
	Male	7	11
	total	36	27
The age range	Minimum date of birth	02.10.2001	04.07.1999
	Maximum date of birth	08.09.2004	20.09.2004
	All students	18 to 21 years	18 to 24 years

Among the 63 participating students, 38 (60.3%) held high school diplomas in the humanities, 13 (20.6%) in experimental sciences, 10 (15.2%) in mathematics and physics, and 2 (3.2%) in Islamic studies. The average high school GPA across both groups was 17.45 (SD = 1.49). The sample also included four guest students from other universities and three students who had transferred into the accounting program from other undergraduate majors. A review of academic records from previous semesters indicated that three students had enrolled in the Principles of Accounting I course but failed to pass.

4.1. Survey about the effectiveness of game-based teaching on the Kahoot platform

H1: Students hold a positive perception regarding their experience with the Kahoot! Platform for teaching Principles of Accounting II.

To examine the perceived impact of using Kahoot! In a game-based accounting classroom environment, students in the experimental group completed a researcher-developed questionnaire. The results are presented in Table 4.

The results indicate that over half of the experimental group students agreed with most questionnaire items, except for item 19, which concerned increased confidence at the end versus learning gains throughout the class.

Items 9 and 10 showed a strong positive attitude (more than 90% in agreement) towards Kahoot! Perceived as effectively motivating cooperative learning and involving fun in-class activities. Responses in favor of above 80% are evident for items 8, 12, 15, 22, and 24. This indicated that students were more engaged and excited when using Kahoot!, less aware of time passing, and more likely to enjoy the learning experience. Additionally, the topics you discuss while playing Kahoot! Sessions promoted more content from an educational perspective, and it was felt that the platform contributed to a better understanding of basic concepts. Items 1, 2, 3, 4, 5, and rotas (item pos: 14; agreement >70%; item pos: 20) demonstrated fair levels of >70%. These responses emphasized Kahoot! As an excellent revision tool before exams. Students were eager to engage with the platform more often in class. They, too, recognized it for improving clarity in particular areas such as inventory, cash, and payroll. Most students found the Kahoot! THOUGHTS: Challenges engaged and stated interest in using it again in future sessions.

Table 4. Results of the Researcher- Made Questionnaire

Items	Percentage of agree and strongly agree	Percentage of disagree and strongly disagree	Indifferent Percentage
The value of using Kahoot in the classroom:			
1. Kahoot! was a good way to review the material before the exam.	72.2%	5.6%	22.2%
2. I recommend that this method be used in future classes as well.	75%	16.7%	8.3%
The effectiveness of Kahoot in understanding the course materials:			
3. The game helped me to better understand the contents of the inventory topic.	75%	8.3%	16.7%
4. The game helped me better understand cash accounting.	75%	8.3%	16.7%
5. The game helped me better understand payroll accounting.	75%	2.8%	22.2%
6. The game helped me better understand the chapter's contents on the partnership (profit sharing and the entry of a new partner).	69.4%	11.1%	19.5%
7. The game helped me to better understand the contents of the chapter on the partnership (New partner entry and exit of the partner in the partnership).	69.4%	5.6%	25%
8. Kahoot meetings helped me learn the topics discussed in the class.	83.3%	2.8%	13.9%
Benefits of Using Kahoot Inside the Classroom:			
9. Kahoot! motivates cooperative learning.	94.4%	0	5.6%
10. Using Kahoot! makes the process of class activities more fun.	94.4%	2.8%	2.8%
11. Immediately after answering the questions, I received information about my success/failure.	69.4%	5.6%	25%
12. Using Kahoot! in the classroom environment and reviewing the material in this way helped me to understand the material.	86.1%	2.8%	11.1%
13. Competition in class gave me a good motivation to be accurate and perform better in learning accounting.	66.7%	8.3%	25%
Elements of Kahoot! Itself:			
14. It was challenging for me to use Kahoot!	77.8%	8.3%	13.9%
15. Using Kahoot! was enjoyable and attractive for me.	88.8%	5.6%	5.6%
16. I felt satisfied with the use of Kahoot during the semester.	66.7%	5.6%	27.8%
17. It was uncomfortable for me when I couldn't connect to Kahoot.	52.8%	19.4%	27.8%
18. Kahoot! was well-designed and organized, and I was able to understand its content and educational goals.	69.5%	8.3%	22.2%
19. At the end of the Kahoot! sessions in class, I was more confident about my learning.	41.7%	22.2%	36.1%
20. I would like to use Kahoot! again.	75%	8.3%	16.7%
21. I did not feel tense or anxious while using Kahoot! In class.	61.1%	13.9%	25%
22. While using Kahoot!, I lost track of time. (class engagement)	86.1%	2.8%	11.1%
23. Using Kahoot! helped me to realize my weaknesses in various topics. (Feedback)	55.6%	8.3%	36.1%
24. My enthusiasm and motivation to learn the concepts of the Principles of Accounting Course II while using Kahoot was greater than that of solving the exercises at home in the traditional and usual way.	86.1%	0	13.9%
25. The level of Kahoot's challenge was proportional to my level of ability.	63.9%	8.3%	27.8%

Six of the items (6, 7, 11, 13, 16, and 9) reported over 60% agreement. These included perceptions of the platform's desired level of challenge, its competitive climate, and its ability to increase accuracy and speed. Kahoot! immediately gave students feedback. This not only enriched their understanding

of course content and learning outcomes in a low-anxiety environment but also bolstered the overall student satisfaction with the semester. The platform was viewed as beneficial for promoting understanding of partnership-related topics.

To test the first hypothesis, a one-sample t-test was conducted. The test values were determined at the midpoint of the 5-point Likert scale (score 3) used in the questionnaire. This method is appropriate when evaluating whether the mean response differs meaningfully from a theoretical reference point rather than comparing two groups. Since each dimension consists of a different number of items, the test value for each dimension was calculated by multiplying the number of items in that dimension by 3 (the midpoint of the scale). Specifically: (1) The value of using Kahoot! In the classroom (2 items): Test value = $2 \times 3 = 6$; (2) the effectiveness of Kahoot! In understanding the course materials (6 items): Test value = $6 \times 3 = 18$; (3) Benefits of using Kahoot! Inside the classroom (5 items): Test value = $5 \times 3 = 15$; (4) Elements of Kahoot! Itself (12 items): Test value = $12 \times 3 = 36$. A mean score significantly higher than the test value indicates a positive perception toward that dimension. As presented in Table 5, all four measured dimensions were statistically significant ($p \leq 0.001$), thus confirming the first research hypothesis. Consequently, it can be concluded that students held a positive perception of their experience with Kahoot! in teaching Principles of Accounting II.

Table 5. One-sample t-test to examine the items

Items	Average	Test Value	t	df	Sig.
The value of using Kahoot! in the classroom	7.800	6	6.120	35	0.001
The effectiveness of Kahoot! in understanding the course materials	22.970	18	8.170	35	0.001
Benefits of using Kahoot! inside the classroom	21.310	15	11.770	35	0.001
Elements of Kahoot! Itself	45.750	36	8.670	35	0.001

Note: Test values represent the midpoint of the 5-point Likert scale (score of 3) multiplied by the number of items in each dimension.

4.2. Investigating the effectiveness of Kahoot! On academic performance

To evaluate the association between Kahoot! Regarding students' academic performance in the Principles of Accounting II course, the Kahoot!-based instructional approach was compared with the traditional teaching method. In this study, students' scores on Principles of Accounting I served as a pre-test (covariate). In contrast, their scores on Principles of Accounting II served as the post-test (dependent variable). To test the second and third hypotheses, both univariate and multivariate analyses of covariance (ANCOVA and MANCOVA) were employed.

Statistical Assumptions and Preliminary Analyses: The underlying statistical assumptions for ANCOVA and MANCOVA were tested before conducting the analyses. Below is the report of the output from these checks against our assumptions. Once these assumptions were confirmed, the principal statistical tests could be performed.

Normality: The Shapiro-Wilk test was used to assess the normality of the dependent variables across groups. The results indicated that the assumption of normality was met for all five classroom tests in both groups ($p > 0.05$ for all variables).

Homogeneity of variance: Levene's test of equality of error variances was conducted for each dependent variable. The results showed that the assumption of equal variances was satisfied for all five tests (Inventory: $F = 1.23$, $p = 0.27$; Cash: $F = 2.01$, $p = 0.16$; Salary: $F = 0.89$, $p = 0.35$; Partnership I: $F = 1.67$, $p = 0.20$; Partnership II: $F = 0.54$, $p = 0.46$).

Homogeneity of regression slopes: The interaction term of each covariate (the pre-test) and independent variable (group) was calculated separately for all dependent variables. The results confirmed that the assumption of homogeneous regression slopes was satisfied for all five tests ($p > 0.05$ for all interactions).

Equality of covariance matrices: Box's test of equality of covariance matrices was conducted to assess the assumption of equal covariance matrices across groups. This yielded a non-significant result (Box's $M = 15.34$, $F = 1.02$, $p = 0.43$), indicating that the covariance matrices were equal between groups and confirming that, despite unequal sample sizes, these are robust to MANCOVA.

MANCOVA was chosen as the main approach to test Hypothesis 3, where all five classroom tests (Inventory, Cash, Salary, Partnership I and II) are conceptually interrelated as measures of formative assessment performance for the same course. MANCOVA enables simultaneous testing of group differences across multiple dependent variables while controlling for the pretest covariate and maintaining the Type I error rate of interest. This method is better than performing several individual ANCOVAs (one for each of the five BTCs), which would increase family-wise error.

H2: There is a statistically significant difference in academic performance in Principles of Accounting II between students who used Kahoot! And those who received traditional instruction.

As shown in Table 6, the experimental group had a higher descriptive mean post-test score (12.31, $SD = 3.44$) compared to the control group (11.83, $SD = 3.20$). However, this descriptive difference does not account for pre-existing differences between the groups. The ANCOVA results in Table 7 show no statistically significant difference between the adjusted mean scores of academic performance in the post-test stage for the experimental and control groups after controlling for the pre-test effect ($F = 0.02$, $p = 0.87$, $\eta_p^2 = 0.001$, Cohen's $d = 0.04$). A nonsignificant ANCOVA indicates that the difference in post-test means was due mainly to differences in pre-test means (rather than the effect of Group). The low observational statistical power for the group effect (0.05) indicates that this analysis would likely not be able to identify small or moderate group differences. Thus, the lack of statistical significance may partly relate to the reduced sample size, as a second randomization did not reflect the complete absence of impact. Consequently, no statistically significant difference in academic performance was established between the two groups, and Hypothesis 2 is rejected.

Table 6. Mean and standard deviation of variables

Variable		Experimental Group		Control Group	
		Mean	SD	Mean	SD
Classroom Tests	Inventory	6.800	5.720	9.920	3.170
	Cash	11.110	6.320	8.560	3.480
	Salary	11.810	4.280	11.190	2.970
	Partnership 1	6.600	3.060	9.600	3.910
	Partnership 2	9.080	3.950	9.680	3.590
Pre-test	Principles of Accounting I	14.800	2.330	14.380	1.830
Post-test	Principles of Accounting II	12.310	3.440	11.830	3.200

H3: There is a statistically significant difference in classroom test performance between students who received Kahoot!-based formative assessment and those who received traditional paper-based formative assessment.

Table 7. ANCOVA results for academic performance (Post-test)

Source of variation	Sum of squares	Degree of Freedom	Mean of Square	F	Sig.	Partial Eta Squared	Cohen's <i>d</i>	Observed Power
Pre-test	218.410	1	218.410	28.240	0.001	0.320	—	0.990
Groups	0.210	1	0.210	0.020	0.870	0.001	0.040	0.050
Error	464.030	60	7.730					
Total	9926.870	63						

The difference between the experimental and control groups lies in how they respond to the functional phase tests. Multivariate analysis of covariance (MANCOVA) was used to investigate this hypothesis. MANCOVA was chosen because the five classroom tests represent related dimensions of student performance, and analyzing them simultaneously controls for intercorrelations among the dependent variables while protecting against inflated Type I error rates that would result from separate univariate tests. The significance values of Pillai's Trace, Wilks' Lambda, Hotelling's Trace, and Roy's Largest Root were examined. The multivariate test results indicated a non-significant overall effect of group across all dependent variables (Pillai's Trace = 0.12, $F = 1.45$, $p = 0.22$, $\eta_p^2 = 0.12$), suggesting that when considered jointly, the five classroom tests did not differ significantly between the experimental and control groups after controlling for pre-test scores.

Table 8. MANCOVA and follow-up univariate ANCOVAs for classroom tests

Table 8: MANCOV and Follow-up Univariate ANCOVAs for Classroom Tests								
Multivariate Tests								
Effect		Pillai's Trace	F	Hypothesis df	Error df	Sig.	Partial Eta Squared	Cohen's <i>d</i>
Univariate Tests	Pre-test	0.080	1.02	5	57	0.410	0.080	
	Group	0.120	1.45	5	57	0.220	0.120	—
Source		Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Cohen's <i>d</i>
Inventory	Pre-test	12.450	1	12.450	1.890	0.170	0.030	
	Group	18.230	1	18.230	2.760	0.100	0.040	-0.670
	Error	396.780	60	6.610				
Cash	Pre-test	8.920	1	8.920	1.340	0.250	0.020	
	Group	15.670	1	15.670	2.350	0.130	0.040	0.490
	Error	399.450	60	6.660				
Salary	Pre-test	3.210	1	3.210	0.480	0.490	0.010	
	Group	1.890	1	1.890	0.280	0.600	0.010	0.160
	Error	401.230	60	6.690				
Partnership I	Pre-test	22.340	1	22.340	3.450	0.070	0.050	
	Group	67.730	1	67.730	5.270	0.020	0.130	-0.870
	Error	464.030	60	7.730				
Partnership II	Pre-test	5.670	1	5.670	0.850	0.360	0.010	
	Group	4.120	1	4.120	0.620	0.430	0.010	-0.160
	Error	398.900	60	6.650				

Note: The multivariate test (omnibus) for the group effect was non-significant ($p = 0.22$). Although the univariate ANCOVA for Partnership I showed a nominal p -value of 0.02, it did not reach the statistical significance threshold after applying the Bonferroni correction ($\alpha = 0.01$). Observed power for the group effect was 0.05.

For a more fine-grained analysis, we performed a one-way ANCOVA for each of the five classroom tests (Inventory, Cash, Salary, Partnership I, and Partnership II), with pre-test scores entered as a covariate. Note that although the overall multivariate test was non-significant, these follow-up univariate tests were thus conducted to further examine this pattern of results across individual tests. This is in line with recommendations on interpreting MANCOVA results, in which univariate analyses can be important for understanding specific variables when the multivariate effect is not significant. Table 8 presents the complete results of these analyses for all five tests.

As shown in the univariate follow-up analyses in Table 8, after controlling for pre-test scores, only the Partnership I test showed a nominally significant univariate difference between the experimental and control groups ($F = 5.27, p = 0.02, \eta_p^2 = 0.13$, Cohen's $d = -0.87$). Given that the overall multivariate test was non-significant, this univariate result should be interpreted with extreme caution. To control for the Type I error rate across the five dependent variables, a Bonferroni correction was applied, resulting in an adjusted significance threshold of 0.01 (0.05/5). Using this more stringent alpha level, the p-value for Partnership I (0.02) is not statistically significant.

Results of the pairwise comparison were shown in Table 9, which showed a mean difference of -2.92 for Partnership I, although since the omnibus test was not statistically significant and since statistical power is very low (0.05) for the group effect when $N=26$, this univariate result could be considered most feasibly as spurious or incidental to a null effect from an intervention. Although students reported enhanced understanding of the material through the use of Kahoot!, their performance on classroom tests did not significantly improve compared with those assessed using traditional paper-based methods. Consequently, the third hypothesis is rejected.

Table 9. Pairwise comparison for Partnership I (Nominal results)

Variable	Group I	Group J	Mean differences (I – J)	Std. Error	Significance level
Partnership I	Experimental	Control	-2.920	1.270	0.020

Note: The reported p-value is nominal. Under the Bonferroni-corrected alpha level ($\alpha = 0.01$ for five comparisons), this difference is not statistically significant.

5. Discussion

This study examined the use of Kahoot!, a mobile-based gamified learning platform, on students enrolled in the Principles of Accounting Course II, focusing on engagement, motivation, and academic performance. Generation Z students, being digital natives, are naturally inclined towards technology-mediated learning environments (Wang and Tahir, 2020). Consistent with previous research in this domain, our analyses show that students who played Kahoot! RESEARCH REPORTS reported increased motivation, engagement, and enjoyment of the classroom (Minton and Bligh, 2021; Marsa et al., 2021). The gamified nature of the platform promotes collaborative learning and instant feedback, creating a feverish educational environment (Neerupa et al., 2024).

Analysis of the First Hypothesis: Our results demonstrate that Kahoot! This is supported by the findings of Lin et al. (2018) and Tóth et al. (2019), who reported that students experienced positive effects on their learning in Principles of Accounting II, primarily as an exam review tool. Students said that interactive formats helped them understand complex topics more clearly, consistent with Khatibi and Badeleh (2021). Kahoot! Studies (Creel et al. 2021). These outcomes align with the engagement-based learning framework and reflect the activating role game-based tools can play in cognitive and affective engagement, as two key mechanisms underpinning this study's theoretical background. Thus, while Minton and Bligh (2021), for instance, found competition to have time-unlimited negative outcomes, most students in this study found it motivating and performance-

enhancing in a context-dependent case. However, the lower level of agreement on the confidence item (Item 19) indicates that positive emotional engagement does not equate to higher levels of perceived learning or confidence, a crucial distinction already reported in previous research, such as [Ranieri et al. \(2021\)](#).

It is important to note that these findings are based on self-reported perceptions, and the test values used in the one-sample t-test were derived by multiplying the midpoint of the 5-point Likert scale (score of 3) by the number of items in each dimension, as detailed in the Results section. This analytical choice is methodologically appropriate for composite Likert constructs; however, such tests rely on the assumption that Likert-type means meaningfully approximate interval-level properties—a limitation that should be considered when interpreting the magnitude of the perceived learning effects. A gamified learning setting motivates students and enhances student enjoyment, as supported by [Felszeghy et al. \(2019\)](#). Technical problems such as unstable internet, as mentioned by [Lin et al. \(2018\)](#), pose challenges. Indeed, in line with [Ranieri et al. \(2021\)](#), theoretical content fared better than practical. In sum, these results add to the existing literature by placing the findings within a theoretical perspective, building on previous empirical research, and demonstrating how Kahoot! Localisation of CS in a university setting. Results demonstrate a localized implementation of CS in an Iranian university context that enhances perceived understanding, student engagement, and enjoyment; however, it also illuminates contextual and infrastructural factors that likely moderate the effectiveness of quality mechanisms.

Second hypothesis discussion: Academic success of the Kahoot! Results suggest that, compared to the control group, the second hypothesis is rejected. Although the descriptive statistics in Table 6 show that the experimental group had a higher mean post-test score (12.31) than the control group (11.83), this apparent difference is not statistically significant after controlling for pre-test scores via ANCOVA ($F = 0.02, p = 0.87$). Crucially, the observed power for this analysis was remarkably low (0.05), and the effect size was negligible (Cohen's $d = 0.06$). These statistical indicators suggest that the study was underpowered to detect effects smaller than extremely large. Thus, the lack of any statistical significance might be due to the final n () being too small rather than in fact a conclusive non-existence of pedagogical effect. Although much of the earlier gamified learning research shows that students are positively inclined toward gamified learning, few studies explicitly examine a quantifiable change in student learning outcomes. The current result reinforces a burgeoning assertion that favorable perceptions do not constitute unequivocal learning benefits, especially when the gamified experience primarily serves as simply a motivational enabler.

Current findings are similar to those of [Devanipour \(2022\)](#), who reported no overall difference in final test scores, despite student enthusiasm in his study. In contrast, studies by [Wardana and Sagoro \(2019\)](#) and [Ortiz-Martínez et al. 2022](#)) found improvements in post-test and exam performance, suggesting that context and implementation strategies may be key determinants. These mixed results suggest that Kahoot!'s effectiveness is closely tied to curricular alignment, depth of content engagement, and the degree of integration with broader learning activities. These findings suggest that Kahoot! It is more linked to influencing learner engagement than to a direct impact on academic delivery outcomes, sometimes only within the wider context of an active and digital learning framework ([Go et al., 2024](#)). Accordingly, the rejection of the second hypothesis does not undermine the pedagogical value of Kahoot!, but instead underscores the need for intentional design when targeting achievement-related outcomes.

Discussion of the third hypothesis: Despite students' favorable perceptions of Kahoot!, the results from formative assessments showed no significant academic differences at the multivariate level (Pillai's Trace, $p = .22$). Although a nominal significance was initially observed in the follow-up univariate analysis for one test (Partnership I; $F = 5.27, p = 0.02$), this finding did not survive the

more conservative Bonferroni correction (adjusted $\alpha = 0.01$). When considering the substantial effect size for this specific module (Cohen's $d = 0.87$) alongside the extremely low observed power (0.05), this isolated result should be interpreted with caution. It is much more likely to represent either a false positive or a chance anomaly than a solid teaching effect, all the more so from relatively small and unbalanced sample sizes. We cautiously interpret the surprising direction of the nominal difference in Partnership I, in which the experimental group underperformed. The Partnership I test was given towards the end of the semester (see Results section), and procedural observations revealed that many students in the experimental group had postponed their preparation until just before the exam. While military service may better explain distance learning, researchers said this is just speculation, lacking supplementary data on study habits given the small sample size.

Accordingly, the third hypothesis is rejected, in line with [Devanipour's \(2022\)](#) findings in Iran but in contrast to [Ortiz-Martínez et al. \(2022\)](#), who reported performance gains in Spain. While previous studies (such as [Wardana and Sagoro, 2019](#); [Ortiz-Martínez et al., 2022](#)) suggest that gamification improves performance, some others (e.g., [Scalan and Huff, 2020](#); [Chan et al., 2021](#); [White et al., 2021](#)) show little to no measurable effects. These inconsistencies across studies suggest that Kahoot! The role of formative assessment or testing to help you learn rather than just testing what you know alone is not enough, however, without robust scaffolding and opportunities for deeper cognitive processing. Moreover, as [Yao et al. \(2024\)](#) noted, Formative assessments might work best for lower-performing students, a trend that could have been diluted in this study, as pointed out (2024). Also, non-instructional variables, such as the Hawthorne effect ([McCambridge et al., 2014](#)), may have moderated learning outcomes. In sum, these studies suggest that the pedagogical benefits of Kahoot! Likely, the effects of immediacy on engagement, attention, and classroom dynamics are more pronounced in these domains than for performance, which continues to depend upon learner characteristics and specific instructional design choices.

Summary of Findings and Implications for Future Research: Overall, the findings indicate that while Kahoot! Positively influenced students' perceptions and classroom engagement, it did not yield statistically significant improvements in academic performance for Hypotheses 2 and 3. The interpretation of these null results is fundamentally constrained by the study's low statistical power (0.05), which indicates a high risk of Type II error. In other words, the study may have lacked the sensitivity to detect small-to-moderate improvements in performance that might have been present in a larger, more robustly powered sample.

Moreover, the absence of assessments targeting deeper learning outcomes such as conceptual understanding or long-term retention further constrains conclusions about the platform's educational impact. Thus, although the present study does not provide evidence of direct academic gains, its findings are inconclusive rather than dismissive of gamification's potential. Future research employing more rigorous designs, larger and randomized samples, and multidimensional measures of learning is needed to more definitively clarify the relationship between gamified tools and student achievement in accounting education.

5.1. Researcher's observations of the two groups

Bloom (1956) emphasizes that both cognitive entry behaviors and affective entry characteristics influence students' academic performance. Affective entry characteristics refer to students' interest, motivation, and attitudes toward learning new material. At the same time, students' interpretations of their past academic successes or failures may also shape their engagement with new course content. Classroom observations indicated that many students in both groups were primarily focused on achieving higher examination scores. The introduction of a game-based learning environment created a learning experience different from traditional paper-based assessments, which appeared to influence

students' classroom reactions and interaction patterns. However, levels of engagement and participation varied among students in both groups.

Students were expected to use appropriate study strategies, conceptual understanding, and problem-solving skills when answering assessment questions. The results confirmed our expectation that Kahoot! might initially be a bit intimidating for students in the experimental group. On the contrary, the students in the experimental group tended to have less structured preparation patterns at first, but they were already familiar with the Kahoot! -based format increased over time. These observations are offered strictly as qualitative insights rather than as empirical results from statistical analyses to contextualize the instructional setting.

6. Conclusion

The results of this research study suggest that Kahoot! These results were not statistically significant in their impact on students' academic gains, but demonstrated improvements in students' perceptions, engagement, and classroom dynamics in the Accounting Principles II course. As shown in this study, academic performance is a multidimensional construct shaped by motivational and attentional factors, as well as students' interactions with learning tools, not just by evaluation scores. The lack of post-test differences after controlling for pre-test scores reinforces the need to use appropriate controls in quasi-experimental designs. It indicates that improvements in perceptions do not always translate into improvements in performance.

Although Kahoot! Has been effectively implemented in various educational contexts worldwide, its impact in this study appears to be shaped by contextual and infrastructural factors. Furthermore, it is critical to acknowledge that the study's final sample size bounds the statistical conclusions ($N = 63$; 36 experimental and 27 control) and a notably low observed power (0.05). Under these conditions, the study was primarily sensitive to very large effects, potentially overlooking smaller but pedagogically meaningful gains. Barriers to digital engagement were reflected in students' experiences of school: inconsistent internet access and uneven uptake of digital tools in teaching practice across schools. Coupled with this, limited statistical sensitivity in detecting achievement measures may account for the additional engagement benefits experienced by the experimental group compared to gains in achievement. In general, Kahoot! Has the capacity to support engagement but appears to need more robust pedagogical integration and a less volatile infrastructure to make a measurable impact on learning outcomes.

7. Resaerch Implications and limitations

Overall, these findings suggest that policymakers and institutional leaders may want to consider supporting the development of faculty capabilities in implementing game-based learning tools, as well as additional investments in digital infrastructure. For instructors, targeted training, along with blended learning modes that balance gamified assessments with more traditional evaluations in pursuit of collaborative learning, may improve overall learning outcomes for students with diverse needs. For such interventions to be effective, institutional commitment through incentives, technical assistance, and the fostering of digital literacy among faculty and students is needed.

Regarding limitations, this study was constrained by its quasi-experimental design with nonequivalent groups. Although ANCOVA was used to statistically control for initial differences, the final sample size of 63 participants, resulting from the exclusion of three cases in the control group to ensure data integrity, may have limited the study's capacity to detect significant differences. This underpowered design (observed power = 0.05) increases the risk of a Type II error, leaving true effects unobserved. The non-significant MANCOVA results also bolster this. However, a nominal difference was noted on one subtest (Partnership I); this difference lost significance after a conservative

Bonferroni correction, suggesting it may be spurious. The reliance on closed-ended Kahoot! Items were also limited in their evaluation of deeper conceptual understanding. Additionally, fluctuations in internet access inhibited students from engaging. Future studies should utilize a priori power analysis to determine optimal sample sizes, randomization as it relates to treatment-receipt effects, larger (and balanced) participant samples, and better methods of addressing attrition to more definitively elucidate the association between gamified tools and accounting learning (especially outcome-based), ultimately aiding student success. Future research should also adopt a mixed-methods approach to investigate qualitative processes of learning alongside quantitative measures of first-contingent sexual partners, and explicitly opt for a narrower, professionally motivated set of dependent variables to reduce the risk of over-testing. Comparative cross-cultural studies could further illuminate how sociocultural and institutional contexts mediate the effectiveness of gamified learning in accounting education.

Acknowledgement

We sincerely thank everyone who contributed to the preparation of this research.

Conflict of Interest

The authors declare no potential conflict of interest regarding the publication of this work. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication or falsification, double publication or submission, and redundancy, have been fully addressed by the authors.

References

1. Aguiar-Castillo, L., Clavijo-Rodríguez, A., Hernández-López, L., De Saa-Pérez, P. and Pérez-Jiménez, R. (2021). Gamification and deep learning approaches in higher education. *Journal of Hospitality, Leisure, Sport & Tourism Education*, 29, A. 100290. <https://doi.org/10.1016/j.jhlste.2020.100290>
2. Almuntsr, N. M., Muhamad, H. B., San, O. T. and Shah, S. M. (2024). Effects of gamification on accounting students' motivation, enjoyment, and academic performance in learning introductory accounting at Libyan universities. *International Journal of Academic Research in Business and Social Sciences*, 14(5), pp. 1-22. <http://dx.doi.org/10.6007/IJARBS/v14-i5/21527>
3. Anwar, Y. and Mulyadi, M. (2025). Exploring gamification approaches in accounting courses: a systematic review. *Journal of Accounting and Finance*, 25(2), pp. 24-33. <https://doi.org/10.33423/jaf.v25i2.7610>
4. Attali, Y. and Van Der Kleij, F. (2017). Effects of feedback elaboration and feedback timing during computer-based practice in mathematics problem solving. *Computers & Education*, 110, pp. 154-169. <https://doi.org/10.1016/j.compedu.2017.03.012>
5. Ayllón, S., Alsina, Á. and Colomer, J. (2019). Teachers' involvement and students' self-efficacy: Keys to achievement in higher education. *PloS One*, 14(5), A. e0216865. <https://doi.org/10.1371/journal.pone.0216865>
6. Bakar, S. Z. A. (2024). Gamification: enhancing financial education and motivation in non-accounting students. *Borneo Engineering & Advanced Multidisciplinary International Journal*, 3(1), pp. 18-23. <https://beam.pmu.edu.my/index.php/beam/article/view/151>
7. Basten, D. (2017). Gamification. *IEEE Software*, 34(5), pp. 76-81. <https://doi.org/10.1109/MS.2017.3571581>
8. Bawa, P. (2019). Using Kahoot to inspire. *Journal of Educational Technology Systems*, 47(3),

- pp. 373-390. <https://doi.org/10.1177/0047239518804173>
9. Bhavani, G., Mehta, A. and Dubey, S. (2020). Literature review: game-based pedagogy in accounting education. *International Journal of Financial Research*, 11(6), p. 165. <https://doi.org/10.5430/ijfr.v11n6p165>
 10. Bloom, B. S. (1956). Taxonomy of educational objectives: Affective domain (Vol. 2). Longmans, Green. London https://books.google.com/books/about/Taxonomy_of_Educational_Objectives_Affec.html?id=1907AAAAIAAJ
 11. Bovill, C. (2020). Co-creation in learning and teaching: The case for a whole-class approach in higher education. *Higher Education*, 79(6), pp. 1023-1037. <https://doi.org/10.1007/s10734-019-00453-w>
 12. Chan, S. H., Song, Q., Trongmateerut, P. and Rivera, L. H. (2021). Will game-based learning enhance performance?. *International Journal of Accounting & Information Management*, 29(4), pp. 651-668. <https://doi.org/10.1108/IJAIM-07-2021-0136>
 13. Chappuis, J., Stiggins, R. J., Chappuis, S. and Arter, J. (2020). Classroom assessment for student learning: Doing it right--using it well (p. 432). New York, NY, USA: Pearson. https://books.google.com/books/about/Classroom_Assessment_for_Student_Learnin.html?id=2YouAAAAQBAJ
 14. Creel, T., Paz, V. and O'Leary, C. (2021). Kahoot! Gamification in an accounting class. *Journal of Higher Education Theory and Practice*, 21(16), pp. 41-46. <https://doi.org/10.33423/jhetp.v21i16.4910>
 15. Devanipour, L. (2022). Improving the learning of cost accounting 1 by using the game in the accounting class. *Applied Research in Management and Accounting*, 7(25), pp. 105-114. (In Persian). <https://joas.ir/user/articles/3832>
 16. Durso, S. D. O., Reginato, L. and Cornacchione, E. (2020). Gamification in accounting and Students' Skillset. *Advances in Scientific and Applied Accounting*, 12(3), pp. 1-21. <https://doi.org/10.14392/ASAA.2019120305>
 17. Elkhamisy, F. A. A. and Wassef, R. M. (2021). Innovating pathology learning via Kahoot! game-based tool: a quantitative study of students' perceptions and academic performance. *Alexandria Journal of Medicine*, 57(1), pp. 215-223. <https://doi.org/10.1080/20905068.2021.1954413>
 18. Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. and Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, 4(3), p. 272-299. <https://doi.org/10.1037/1082-989X.4.3.272>
 19. Felszeghy, S., Pasonen-Seppänen, S., Koskela, A., Nieminen, P., Härkönen, K., Paldanius, K. M., ... and Mahonen, A. (2019). Using online game-based platforms to improve student performance and engagement in histology teaching. *BMC Medical Education*, 19(1), p. 273. <https://doi.org/10.1186/s12909-019-1701-0>
 20. Fortuna, J. M., de la Fuente, G. and Velasco, P. (2023). Does gamification mediate the relationship between digital social capital and student Performance? A survey-based study in Spain. *The International Journal of Management Education*, 21(3), A. 100846. <https://doi.org/10.1016/j.ijme.2023.100846>
 21. Go, N. M., Riantono, I. E. and Cassandra, C. (2024). Gamification as a form of active and digital learning in accounting education (2019–2024). In *2024 International Conference on ICT for Smart Society (ICISS)* (pp. 1-8). IEEE, New York. <https://doi.org/10.1109/iciss62896.2024.10751061>
 22. Göksün, D. O. and Gürsoy, G. (2019). Comparing success and engagement in gamified learning experiences via Kahoot and Quizizz. *Computers & Education*, 135, pp. 15-29.

- <https://doi.org/10.1016/j.compedu.2019.02.015>
23. Gómez, J. L. and Monroy, L. D. (2018). Gamification in accounting distance education. *Journal of International Scientific Publications*, 12(8), pp. 208-215.
 24. Haji Moradkhani, H., Mashayekh, S. and Rahman, A. (2017). Traditional teaching methods vs. teaching through a game board in the Principles of Accounting (1) course. *Journal of Accounting Advances*, 9(1), pp. 1-30 (In Persian). <https://doi.org/10.22099/jaa.2017.24509.1508>
 25. Ionescu-Feleagă, L., Dragomir, V. D., Rîndașu, S. M., Stoica, O. C., Curea, Ș. C., Bunea, M. and Barna, L. E. L. (2025). Business simulation games from the perspective of accounting and management professors: Implications for sustainability education in universities. *The International Journal of Management Education*, 23(2), A. 101147. <https://doi.org/10.1016/j.ijme.2025.101147>
 26. Jääskä, E. and Aaltonen, K. (2022). Teachers' experiences of using game-based learning methods in project management higher education. *Project Leadership and Society*, 3(1), A. 100041. <https://doi.org/10.1016/j.plas.2022.100041>
 27. Khatibi, M. B. and Badeleh, A. (2021). The potential of using Kahoot to empower EFL learners: An experience of game-based learning. *Research in English Language Pedagogy (RELP)*, 9(3), pp. 44. <https://doi.org/10.30486/relp.2021.1930824.1273>
 28. Krasodomska, J. and Godawska, J. (2021). E-learning in accounting education: The influence of students' characteristics on their engagement and performance. *Accounting Education*, 30(1), pp. 22-41. <https://doi.org/10.1080/09639284.2020.1867874>
 29. Lee, C. C., Hao, Y., Lee, K. S., Sim, S. C. and Huang, C. C. (2019). Investigation of the effects of an online instant response system on students in a middle school of a rural area. *Computers in Human Behavior*, 95, pp. 217-223. <https://doi.org/10.1016/j.chb.2018.11.034>
 30. Lew, C. and Saville, A. (2021). Game-based learning: Teaching principles of economics and investment finance through Monopoly. *The International Journal of Management Education*, 19(3), A. 100567. <https://doi.org/10.1016/j.ijme.2021.100567>
 31. Lin, D. T. A., Ganapathy, M. and Kaur, M. (2018). Kahoot! It: Gamification in higher education. *Pertanika Journal of Social Sciences and Humanities*, 26(1), pp. 565-582. <https://www.researchgate.net/publication/320182671>
 32. Locke, E. A. and Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, 57(9), pp. 705-717. <https://doi.org/10.1037/0003-066X.57.9.705>
 33. López-Hernández, C., Lizarraga-Álvarez, G. I. and Soto-Pérez, M. (2023). Enhancing learning of accounting principles through experiential learning in a board game. *Accounting Education*, 32(3), pp. 300-331. <https://doi.org/10.1080/09639284.2022.2059770>
 34. Marsa, S. S., Kuspiyah, H. R. and Agustina, E. (2021). The effect of Kahoot! game in teaching reading comprehension achievement. *JET (Journal of English Teaching)*, 7(2), pp. 133-149. <https://doi.org/10.33541/jet.v7i2.2738>
 35. McCambridge, J., Witton, J. and Elbourne, D. R. (2014). Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *Journal of Clinical Epidemiology*, 67(3), pp. 267-277. <https://doi.org/10.1016/j.jclinepi.2013.08.015>
 36. Mekler, E. D., Brühlmann, F., Opwis, K. and Tuch, A. N. (2013). Disassembling gamification: the effects of points and meaning on user motivation and performance. In *CHI'13 extended abstracts on human factors in computing systems* (pp. 1137-1142). Paris, France. <https://edoc.unibas.ch/entities/publication/8a1f71da-d234-42eb-9110-b658c6e1efd5>
 37. Minton, M. and Bligh, B. (2021). Examining the use of Kahoot to support digital game-based

- formative assessments in UAE higher education. *Studies in Technology Enhanced Learning*, 1(2), pp. 445-462. <https://doi.org/10.21428/8c225f6e.32b8666f>
38. Montola, M., Stenros, J., Waern, A. (2009). *Pervasive Games. Experiences on the Boundary between Life and Play*. Taylor & Francis Group, Boca Raton, Florida. <https://doi.org/10.1201/9780080889795>
 39. Neerupa, C., Naveen Kumar, R., Pavithra, R. and John William, A. (2024). Game on for learning: a holistic exploration of gamification's impact on student engagement and academic performance in educational environments. *Management Matters*, 21(1), pp. 38-53. <https://doi.org/10.1108/MANM-01-2024-0001>
 40. Nkhoma, C., Nkhoma, M., Thomas, S., Tu, L. K. and Le, N. Q. (2018). Gamifying a flipped first year accounting classroom using Kahoot. *International Journal of Information System and Engineering*, 6(2), 93-115. <https://dx.doi.org/10.24924/ijise/2018.11/v6.iss2/93.115>
 41. Omar, N., Mohd Isa, N. S., Halid, S., Sulaiman, N., Salleh, N. A. and Abdul, R. (2024). Investigating the impact of gamification to encourage learning motivation among undergraduate accounting students. In *e-Proceedings International Conference on Accounting & Business (ICAB) 2024* (pp. 163-169). <https://ir.uitm.edu.my/id/eprint/134078>
 42. Ortiz-Martínez, E., Santos-Jaen, J. M. and Palacios-Manzano, M. (2022). Games in the classroom? Analysis of their effects on financial accounting marks in higher education. *The International Journal of Management Education*, 20(1), A. 100584. <https://doi.org/10.1016/j.ijme.2021.100584>
 43. Özdemir, O. (2025). Kahoot! Game-based digital learning platform: A comprehensive meta-analysis. *Journal of Computer Assisted Learning*, 41(1), A. e13084. <https://doi.org/10.1111/jcal.13084>
 44. Pitic, D. and Irimias, T. (2023). Enhancing students' engagement through a business simulation game: A qualitative study within a higher education management course. *The International Journal of Management Education*, 21(3), A. 100839. <https://doi.org/10.1016/j.ijme.2023.100839>
 45. Queiro Ameijeiras, C. M., Marti Parreño, J. and Seguí Mas, E. (2025). Beyond the ledger: A review of gamification in accounting education (2014-2023). *Intangible Capital*, 21(3), pp. 508-533. <https://doi.org/10.3926/ic.3139>
 46. Ramesh, G., Azzah, A. M., Habiba, M. S. A. M. and Pauline, S. M. (2025). Gamification and students' engagement in accounting courses—an experimental study. *Learning and Teaching in Higher Education: Gulf perspectives*, 19(1), pp. 35-48. <https://doi.org/10.1108/LTHE-12-2020-0063>
 47. Ranieri, M., Raffaghelli, J. E. and Bruni, I. (2021). Game-based student response system: Revisiting its potentials and criticalities in large-size classes. *Active Learning in Higher Education*, 22(2), pp. 129-142. <https://doi.org/10.1177/1469787418812667>
 48. Sailer, M. and Homner, L. (2020). The gamification of learning: A meta-analysis. *Educational Psychology Review*, 32(1), pp. 77-112. <https://doi.org/10.1007/S10648-019-09498-W>
 49. Scalan, G. and Huff, K. (2020). Incorporating innovative pedagogies into accounting curriculum: results from a game lab implementation. *Journal of Higher Education Theory and Practice*, 20(1), pp. 141-150. <https://doi.org/10.33423/jhetp.v20i1.2785>
 50. Senali, M. G., Iranmanesh, M., Ghobakhloo, M., Gengatharen, D., Tseng, M. L. and Nilsashi, M. (2022). Flipped classroom in business and entrepreneurship education: A systematic review and future research agenda. *The International Journal of Management Education*, 20(1), A. 100614. <https://doi.org/10.1016/j.ijme.2022.100614>
 51. Silva, R., Rodrigues, R. and Leal, C. (2021). Games based learning in accounting education –

- which dimensions are the most relevant? *Accounting Education*, 30(2), pp. 159-187. <https://doi.org/10.1080/09639284.2021.1891107>
52. Silva, S., Rodrigues, R. and Leal, C. (2019). Play it again: how game-based learning improves flow in Accounting and Marketing education, *Accounting Education*, 28(3), pp. 1-24. <https://doi.org/10.1080/09639284.2019.1647859>
 53. Tóth, Á., Lógó, P. and Lógó, E. (2019). The effect of the Kahoot quiz on the student's results in the exam. *Periodica Polytechnica Social and Management Sciences*, 27(2), pp. 173-179. <https://doi.org/10.3311/PPso.12464>
 54. Turan, Z. and Meral, E. (2018). Game-based versus to non-game-based: the impact of student response systems on students' achievements, engagement, and test anxieties. *Informatics in Education*, 17(1), pp. 105-116. <https://doi.org/10.15388/infedu.2018.07>
 55. Vizcaya-Moreno, M. F. and Pérez-Cañaveras, R. M. (2020). Social media used and teaching methods preferred by generation z students in the nursing clinical learning environment: A cross-sectional research study. *International Journal of Environmental Research and Public Health*, 17(21), A. 8267. <https://doi.org/10.3390/ijerph17218267>
 56. Vroom, V. H. (1964). *Work and motivation*. Wiley. Montreal, Quebec, Canada
 57. Wang, A. I. and Tahir, R. (2020). The effect of using Kahoot! for learning—A literature review. *Computers & Education*, 149, A. 103818. <https://doi.org/10.1016/j.compedu.2020.103818>
 58. Wang, A. I., Zhu, M. and Sætre, R. (2016). The effect of digitizing and gamifying quizzing in classrooms. In *Proceedings of the 10th European Conference on Games Based Learning* (pp. 729-737). Sonning common: academic conferences and publishing international, Trondheim, Norway. <http://hdl.handle.net/11250/2426374>
 59. Wanner, T. and Palmer, E. (2018). Formative self-and peer assessment for improved student learning: the crucial factors of design, teacher participation and feedback. *Assessment & Evaluation in Higher Education*, 43(7), pp. 1032-1047. <https://doi.org/10.1080/02602938.2018.1427698>
 60. Wardana, S. and Sagoro, E. M. (2019). Implementasi gamifikasi berbantu media kahoot untuk meningkatkan aktivitas belajar, motivasi belajar, dan hasil belajar jurnal penyesuaian siswa kelas X akuntansi 3 di SMK Koperasi Yogyakarta tahun ajaran 2018/2019. *Jurnal pendidikan akuntansi indonesia*, 17(2), pp. 46-57. <https://doi.org/10.21831/jpai.v17i2.28693>
 61. White, T., Brody, R. G. and Gupta, G. (2021). Who's got next? An analysis of the inhibitors to mobile game adoption in an introductory accounting class. *Advances in Accounting Education: Teaching and Curriculum Innovations*, 25, pp. 23-47. <https://doi.org/10.1108/S1085-462220210000025002>
 62. Yao, Y., Amos, M., Snider, K. and Brown, T. (2024). The impact of formative assessment on K-12 learning: A meta-analysis. *Educational Research and Evaluation*, 29(7-8), pp. 452-475. <https://doi.org/10.1080/13803611.2024.2363831>
 63. Zhang, P. (2008). Motivational affordances: Reasons for ICT design and use. *Communications of the ACM*, 51(11), pp. 145-147. <https://doi.org/10.1145/1400214.1400244>



RESEARCH ARTICLE

The Moderating Effect of Power Distance on the Relationship between Conscientiousness and Extraversion with Auditors' Impartiality

Hamid Alvari Chenari, Omid Pourheidari*, Ahmad Khodamipour

Department of Accounting, Faculty of Management and Economics, Shahid Bahonar University of Kerman, Kerman, Iran

How to cite this article:

Alvari Chenari, H. , Poorheidari, O. and Khodamipour, A. (2026). The Moderating Effect of Power Distance on the Relationship between Conscientiousness and Extraversion with Auditors' Impartiality. Iranian Journal of Accounting, Auditing and Finance, 10(3), 139-161. doi: 10.22067/ijaaf.2026.48274.1644
https://ijaaf.um.ac.ir/article_48274.html

ARTICLE INFO

Article History

Received: 2025-08-27

Accepted: 2025-12-02

Published online: 2026-07-12

Keywords:

Auditors' Impartiality,
Auditors' Conscientiousness,
Auditors' Extraversion,
Auditors' Independence, Power
Distance

Abstract

Auditors' impartiality is fundamental for responsible and logically grounded audit opinions. While prior research has examined the link between personality characteristics (conscientiousness and extraversion) and professional judgment, the moderating role of power distance remains largely unexplored in the auditing context. Existing studies have focused on structural aspects of independence, overlooking psychological and cultural contingencies. This study is the first to introduce power distance as a moderating variable between the Big Five personality characteristics and auditors' impartiality, adopting a contingency perspective that challenges the one-size-fits-all assumption in the auditing behavior literature. The present study is applied in nature regarding its purpose and methodology, specifically utilizing structural equation modeling. The data collection tool was a questionnaire. The study population comprised auditors at all levels within auditing firms, and the sample size was 384 auditors. These findings showed that the study's independent variables, conscientiousness and extraversion, have a positive and significant effect on impartiality. However, power distance moderates only the relationship between extraversion and impartiality, not that between conscientiousness and impartiality. Therefore, the power distance variable has a significant moderating effect on only one of the auditors' personality dimensions, namely extraversion. This highlights the importance of a contingency approach to understanding cultural impact, recognizing that cultural factors do not uniformly influence all individual characteristics but tend to have a stronger negative effect on those related to asserting or challenging authority, such as extraversion.

<https://doi.org/10.22067/ijaaf.2026.48274.1644>



NUMBER OF REFERENCES
60



NUMBER OF FIGURES
4



NUMBER OF TABLES
12

Homepage: <https://ijaaf.um.ac.ir>
E-ISSN: 2717-4131
P-ISSN: 2588-6142

*Corresponding Author: Omid Pourheidari
Email: opourheidari@uk.ac.ir
Tel:
ORCID: 0000-0003-1599-4434

1. Introduction

Auditor impartiality is defined by the professional code of conduct: a certified public accountant must not allow bias, conflicts of interest, or undue influence by others to affect their professional judgment or work (IESBA, 2018). The principle of impartiality requires all certified public accountants to avoid bias, conflicts of interest, and undue influence from others in their professional judgments or business activities. Accordingly, this study defines auditor impartiality as the auditor's commitment to making professional judgments based solely on audit evidence and applicable standards, without favoritism toward any stakeholder (e.g., management, shareholders, or regulatory bodies). Impartiality is distinct from independence in that it is a behavioral and cognitive condition reflecting the extent to which the auditor's decision-making process appears unaffected by conflicting relationships. In addition, impartiality is the state of being unbiased towards the interests of any party, whereas objectivity relies on fact-based evidence not open to interpretation. Thus, impartiality is a separate construct, conceptually situated between structural independence and objective judgment, as it endeavors to describe an auditor's behavior when faced with competing interests.

Auditors' impartiality is an essential characteristic that enables them to provide responsible opinions grounded in a sound and logical understanding of professional conduct (Nelson, 2009). Because managers have discretion in applying accounting standards and procedures, auditors must rely on professional judgment and maintain impartiality when expressing their opinions. An essential requirement of impartiality is that auditors' judgments and decisions are not subordinate to others (Svanberg and Öhman, 2016). However, auditors' impartiality is influenced by the work environment, relationships with management, incentives received, client recognition, abilities, and various personal and organizational characteristics (Samagaio and Felício, 2022). The interaction between personality characteristics and ethical commitment suggests that auditors with specific characteristics are more likely to maintain impartiality, thereby improving audit quality (Haghighi et al., 2024).

Particular traits of an auditor, such as conscientiousness and extraversion, are positively correlated with professional skepticism, which is important to maintain neutrality in the field of auditing. Conversely, neuroticism is associated with reduced audit quality, meaning that high levels of certain personality traits can affect impartiality (Wijaya and Cheisviyanny, 2024). Lack of supervision and time constraints may lead to behaviors such as cyberloafing, which reduce personal gain and negatively impact professional performance (Hessari et al., 2025).

Various factors that impact auditor impartiality also include personality characteristics and cultural dimensions (see Khurana et al., 2026; Veledar et al. 2026). However, no study has investigated the moderating effect of power distance on the relationship between personality traits (i.e., conscientiousness and extraversion) and auditor independence among auditors in the Iranian context. In the Iranian context, this gap is particularly important due to high power distance (Hofstede, 1984). In this regard, the present study aims to answer the following research questions:

RQ1: To what extent do conscientiousness and extraversion significantly affect auditors' impartiality?

RQ2: How does power distance moderate the relationship between each personality trait (i.e., conscientiousness and extraversion) and auditors' impartiality?

RQ3: Does the moderating effect of power distance differ between conscientiousness and extraversion?

The primary objective of this study is, therefore, to examine the moderating effect of power distance on the relationship among conscientiousness, extraversion, and auditor impartiality.

Trust in auditors would enhance auditors' performance and, in turn, increase audit quality; organizational culture is critical to developing trust within audit firms. However, when the distance of power widens through greater separation and limited voice among managers, partners, supervisors,

and auditors within organizations, it can pose a significant risk to auditor independence, particularly in Iranian audits.

This study makes several notable contributions. First, it bridges the gap among personality psychology, auditing, and cross-cultural management by introducing power distance as a moderating variable in the relationship between personality characteristics (conscientiousness and extraversion) and auditor impartiality. This link has not been examined previously. Second, it offers a fresh perspective on auditor impartiality by conceptualizing it as a distinct behavioral and cognitive construct, separate from independence and objectivity. Third, it provides empirical evidence from a high-power-distance context (Iran), thereby enriching the understanding of how cultural dimensions influence professional judgment in developing countries. In practice, the findings can inform audit firms' recruitment, training, and performance evaluation and assist regulators in developing ethical standards tailored to hierarchical environments. Academics and researchers may also benefit from the theoretical framework proposed here as a basis for future cross-cultural auditing research.

In the next section, we will review the relevant literature, theoretical foundations, and background related to the topic of this research. The third section will outline the research methodology. Following that, the fourth section will present the research findings. Finally, the fifth section will present the study's conclusion.

2. Theoretical foundations and development of hypotheses

2.1. Theoretical issues

This study uses the Big Five personality model (Costa and McCrae, 1992), which is one of the best-known and empirically appropriate models for understanding persons. The theoretical aspect of personality: the Big Five model identifies five broad dimensions, including Conscientiousness. Extraversion. Neuroticism. Openness to experience. The Agreeableness and Conscientiousness indices include characteristics such as responsibility, diligence, orderliness, and adherence to standards; Extraversion includes assertiveness and sociability (Costa and McCrae, 1992). This two-dimensional classification of issues is relevant to the watering down of standards (whether auditors should accept audit opinions that do not conform to standards) and to whether auditors willingly express dissenting views.

According to Self-Determination Theory (SDT), all human behavior is fundamentally driven by the needs for competence, relatedness, and autonomy, with intrinsic motivation serving as a primary driver of behavior (Ryan and Deci, 2000). When these needs are met, individuals are more motivated to perform activities, even without external rewards, and experience better psychological well-being and performance.

Competence is one of the three needs in SDT and, as such, is tied to conscientiousness. Conscientiousness: They find it essential to execute tasks well and fulfill professional standards. This drive aligns with the need for competence in SDT. On the flip side, conscientious auditors may become less motivated to uphold impartial standards when hierarchical pressures challenge their competence)

The link between SDT and extraversion lies in the need for autonomy and relatedness. Extraverted auditors seek social engagement and are sensitive to relational cues. In low-power-distance environments that support autonomy and open communication, extraverted auditors are more likely to express dissenting views and assert their impartial judgments. However, in high-power-distance environments that suppress autonomy, extraverted auditors may internalize relational pressures, leading to self-censorship and reduced impartiality.

In high-power-distance cultures, the need for autonomy is often suppressed because individuals

are expected to defer to superiors. Consequently, auditors' intrinsic motivation to uphold impartiality, which requires autonomy, is weakened. In contrast, low-power-distance cultures promote independence, strengthening auditors' intrinsic motivation to maintain impartiality.

The Cognitive Dissonance Theory (CDT) argues that the human mind holds multiple opposing cognitions, so they do not reach equilibrium, resulting in discordant behavior or thoughts, leading to negative feelings and forcing the person to resolve it by changing their cognitions (Festinger, 1957). The nature of being a professional auditor entails a sincere belief in the principles of neutrality. Causal environment: A situation in which an auditor is compelled, or elects to act in a way that undermines their impartiality (e.g., because of client or supervisor pressure) creates cognitive dissonance.

Conscientious individuals have stronger internal ethical standards and are therefore more likely to experience greater initial dissonance when their behavior violates impartiality norms. To reduce this discomfort, they may adopt different strategies depending on power distance: (a) in high-power-distance cultures, they may reduce dissonance by justifying obedience to authority (e.g., "I must obey my superiors"); (b) in low-power-distance cultures, they may change their behavior to restore alignment with their values.

Extraverted individuals, being more socially oriented, may seek external validation to reduce dissonance. They may consult with colleagues or seek social approval for their behavior to minimize incongruence. In high-power-distance cultures, this may lead them to conform to hierarchical expectations; in low-power-distance cultures, they may seek peer support for impartial actions.

In high-power-distance cultures, the dissonance resulting from a violation of impartiality may be reduced more quickly through ready-made justifications such as "I must obey my superiors" or "My position does not permit dissent." That is, the psychological cost of incongruence is less keenly felt because society provides ready-made justifications for behavior that is inconsistent with impartiality. In low-power-distance cultures, dissonance is felt more strongly due to values of autonomy and individual responsibility, and the individual may feel greater pressure to defend their impartiality.

2.2. Conscientiousness and Impartiality

The Five-Factor Model (FFM) of personality designates conscientiousness as one of the major personality dimensions (Costa and McCrae, 1992). This includes traits such as responsibility, dependability, high-quality workplace performance, hard work, and rule-following (Steel et al., 2008). Conscientious individuals aim to perform their tasks well and are willing to spend more time to achieve their objectives (Jin, 2024). They are characterized by perfectionism and attention to detail, which are associated with remarkable accomplishments in their field (Tiliopoulos et al., 2010).

Conscientiousness can promote auditor impartiality through three interrelated mechanisms:

Pathway 1: Adherence to Professional Standards. Conscientious individuals have a strong internal drive to follow rules and complete tasks thoroughly. In the auditing context, this translates into rigorous adherence to auditing standards and ethical codes (Samagaio and Felício, 2022). When internal standards of duty bind auditors, they are less likely to deviate from impartial judgments, even under pressure.

Pathway 2: Diligent Evidence Evaluation. Conscientiousness positively affects auditors' professional skepticism, leading to more complete and objective evaluations of audit evidence (Wijaya and Cheisviyanny, 2024; Saraswati and Latrini, 2023). Rather than accepting management's assertions at face value, conscientious auditors systematically question and verify information, a core component of impartiality.

Pathway 3: Resistance to Biased Shortcuts. Highly conscientious individuals are less prone to cognitive biases and judgment shortcuts (Tiliopoulos et al., 2010). They invest greater cognitive effort

in decision-making, reducing the likelihood of favoritism or premature conclusions. This cognitive diligence directly supports impartial judgment.

There is also evidence of a positive association between conscientiousness and impartiality-based outcomes. [Gold et al \(2009\)](#) Conscientiousness has been found to enhance auditors' objectivity. [Samagaio and Felício \(2022\)](#) have also found to increase auditor capabilities and professional efficiency, thereby further strengthening immunity against bias.

We argue that conscientiousness encourages auditors to work more diligently and leads them towards greater impartiality. Furthermore, considering the definitions, research findings, and theoretical foundations, we believe that auditor conscientiousness can have a positive and significant effect on auditor impartiality; consequently, we propose the following hypothesis.

H1: Auditors' conscientiousness has a significant positive effect on auditor impartiality.

2.3. Extraversion and Impartiality

People have two critical personality dimensions called extraversion and introversion. These two dimensions are not separate; all people in society have them simultaneously. They can be identified by stating the characteristics of extroverts ([Liegl and Furtner, 2023](#)). Extraverts are typically characterized by a preference for collaborative work over solitary tasks, demonstrating frankness in speech, elevated sociability, and practicality. Furthermore, they exhibit a strong desire for stimulation and excitement, are notably gregarious, and possess considerable confidence in their social interactions ([Wang et al., 2025](#)). Concurrently, extroverts possess positive emotional styles, engage in high levels of social interaction and activity, and maintain very intimate interpersonal relationships ([Bowling et al., 2024](#)).

This is clearly evidenced by the finding that extroverted auditors have more professional skepticism than introverted auditors ([Noegroho et al., 2025](#)). Personality Features: The effects of personality characteristics on the judgment and decision-making process will also directly or indirectly affect fraud detection by auditors ([Suryandari et al., 2026](#)). More hopeful about the future are extroverted auditors, who are more satisfied with their personal relationships. Moreover, they achieve higher positions on the career ladder, earn higher incomes and benefits, and experience greater subjective and internal well-being ([Baral, 2024](#)). They have a greater tendency to question underlying assumptions and proactively demand additional information necessary to conduct an impartial audit, and they perform better in managing perceived audit risks ([Chen et al., 2023](#)). The outlook and positive emotions of extroverted auditors who adequately share their knowledge with group members facilitate the initiation and continuity of group formation ([Anwar, 2017](#)). According to [Kerckhofs et al. \(2024\)](#), highly managed female extroverted auditors have less conflict with their managers. By integrating a series of spirituality indices and the moral atmosphere in the Iranian work environment, the findings show that individual personality dimensions and personality type have a greater effect than any other variable on audit quality ([Nasirpour et al., 2023](#)).

Based on the definitions provided, previous research findings, and the established theoretical framework, we assert that an auditor's extraversion strengthens their capabilities in both internal and external communication, heightens their cognitive awareness, and influences subsequent decision-making; all of which contribute to greater impartiality in their judgments. Consequently, we propose the following hypothesis regarding the positive and significant relationship between auditor extraversion and impartiality:

H2: Auditors' extroversion has a significant positive effect on auditor impartiality.

2.4. Power distance and impartiality

Culture manifests as a system of shared social behaviors developed through interactions among members of a specific social group (Laland and Hoppitt, 2003). Cultural values encompass the collective attitudes, beliefs, and behaviors characteristic of individuals within a specific culture (Schwartz, 2012). Moreover, culture is conceptualized across various dimensions, with power distance being one of the essential constructs (Hofstede, 1984). Power distance, when defined, refers to how much followers/subordinates accept the organizational hierarchy, as well as the level of distance between managers and workers. This particular idea refers to the level of acceptance of unequal balance of power in the organizational context (Hofstede, 1980). Perhaps the most defining feature of low power-distance climates is also one that makes managers most uncomfortable: subordinates in low power-distance environments most often have open channels for communicating their complaints to management, including critical feedback and criticism (Kosłowski et al., 2011). They are more involved in decision-making, face less pressure, and receive greater social support (Qian et al., 2014).

Audit culture is defined as the process through which scientific and technical foundations evolve from traditional to modern states. According to the theoretical literature, developing countries, such as Iran, face a long path to developing an audit culture. Such a pathway would require an enduring cultural emphasis on professional independence and neutrality over hierarchical obedience (Tarighi et al., 2022). Leaders in Iranian organizational culture have learned from experience that openly expressing dissent can have unfortunate personal repercussions. They remain trained in 'saving face' and conforming, which piles up their manpower. This phenomenon is rooted in the inefficient hierarchical structures still prevalent in Iranian administrative and educational institutions (Asnaashari et al., 2025).

The impact of power distance on auditor impartiality is significant, as cultural dimensions shape auditors' perceptions and decision-making. Auditors' perceptions of internal controls and fraud risk are influenced by their artistic knowledge, especially in high-power-distance cultures where hierarchical structures are more prominent (Ríos Figueroa, 2013). Such environments may also lead auditors to be less skeptical of management, thereby increasing the threat to their objectivity. In contrast, others claim a strong corporate culture can foster ethical behavior and mitigate stakeholders' malfeasance, leading to the conclusion that cultural dimensions may not always lead to auditor partiality (Zaman, 2024). In addition, high power distance can be a big hindrance to organizational change and development because it is an important limiting condition of organizational flexibility, and then with the ambition to induce complications such as job dissatisfaction, motivation to change jobs rises (especially among Junior professionals seeking upward mobility), and erosion of social capital (Huang and Van De Vliert, 2003).

Power distance influences auditor impartiality through several mechanisms: (1) discouraging upward communication and critical feedback; (2) reducing professional skepticism through uncritical acceptance of managerial authority; (3) suppressing autonomy, which reduces intrinsic motivation to remain impartial; (4) facilitating cognitive dissonance reduction; and (5) increasing the perceived personal costs of dissent.

A study by Shahzad et al. (2025) examined how despotic leadership affects job performance, job satisfaction, and turnover intention at the employee level, with individual-level power distance orientation as a moderator. Based on multisource, time-lagged survey data from 65 leaders and 300 subordinates, they found that despotic leadership is related to poorer job performance and job satisfaction, but to higher turnover intention. Additionally, both power distance orientation and moderated these relations, such that low power distance employees reacted more strongly to despotic leadership. In line with this, studies have shown that not only demographic characteristics but also

cultural dimensions, such as power distance, play an important role in auditors' professional judgment and decision-making (Nolder and Riley, 2014; Huang and Van de Vliert, 2003).

The findings of Bik and Hooghiemstra (2018) indicated that auditors' perspectives on evaluating internal controls differ significantly depending on whether companies operate in low- or high-power-distance cultural contexts. The survey by Al-Asmakh et al. (2024) examined the effect of cultural dimensions, particularly power distance, on auditors' judgment and decision-making. In a structural equation modeling analysis, the results showed that power distance has a significant negative effect on auditors' judgment and decision-making.

We argue that reducing power distance in audit firms increases their impartiality in judgments and decisions. Based on the clearly established definitions, past research findings, and theoretical rationale, we propose that auditors working within highly power-distance cultures face more obstacles to upholding their independence and to eroding it. For this reason, we suggest that power distance in audit firms has a strong negative impact on auditor independence, which gives rise to our hypothesis:

H3: Power distance has a significant negative effect on auditor impartiality.

2.5. The moderating role of power distance

Power distance reflects the extent to which individuals accept and expect unequal power distribution in organizations and societies (Hofstede, 1980). In the auditing context, power distance shapes hierarchical relationships among auditors, managers, and partners and may discourage or encourage auditors from challenging authority (Khurana et al., 2026).

Though both conscientiousness and extraversion could predictably (see hypothesis 2) interact with power distance (conscientiousness X power-distance; extraversion X power-distance), we argue that these personality traits are not uniformly subject to cultural moderation. Conscientiousness, based on internalized demands for duty, order, and personal responsibility, may be less susceptible to contextual effects such as power distance. Conscientious individuals may rely more on their own internal moral compass, which is less susceptible to hierarchical pressures (Winterich and Zhang, 2014). This suggests that extraversion defined by engagement and festivity, along with other social-process forms of extraversion assertiveness (Wang et al., 2025), sociability, sensitivity to positive social rewards and others' positive or negative behavior might be more open to modification through hierarchical cultural niceties.

In high-power-distance cultures, the diminished sense of personal responsibility (Winterich and Zhang, 2014) might theoretically weaken the positive effect of conscientiousness on impartiality. However, empirical evidence suggests that conscientious individuals tend to maintain their internal ethical standards regardless of hierarchical context (Verma et al., 2025). This resilience may render the conscientiousness-impartiality relationship relatively invariant to power distance. Accordingly, while a theoretical case for moderation can be made, the empirical expectation remains ambiguous.

Extraverts desire social interaction and are attuned to signs of hierarchy (Dust et al., 2021). Extraverts may refrain from assertive behaviors, such as speaking up during professional disagreements (Rhodes et al., 2021), in high-power-distance contexts because voice challenges authority. This can make them less objective when hierarchical pressure conflicts with professional standards. Furthermore, greater power distance reduces the perceived duty to resist organizational pressure (Winterich and Zhang, 2014), which in turn diminishes extraverts' tendency to maintain unbiased judgments. This suggests a greater moderating effect of power distance on extraversion than on conscientiousness.

Direct evidence from accounting research demonstrates that power distance is an important antecedent of ethical decision making and plays a partial mediating role between country of origin

and ethical behavioral intention (Curtis et al., 2012). Similarly, Khurana et al. (2026) show that cultural trust, which is connected to power distance, affects audit partners' decisions (client acceptance/rejection). Findings can be interpreted as evidence for the potential moderating role of power distance expectations in achieving auditor neutrality.

Using previous definitions, real-world empirical results, and theoretical bases for the measures identified above, we hypothesize that power distance moderates the effects of auditors' cognitive conscientiousness and extraversion on their impartiality (H1).

H4: Power distance moderates the relationship between auditors' conscientiousness and impartiality.

H5: Power distance moderates the relationship between auditors' extraversion and impartiality.

3. Methodology

3.1. Research method

Since this study seeks to explain the effects of auditors' conscientiousness and extraversion on auditors' impartiality, and to examine the moderating role of power distance in this relationship, this research is a descriptive survey. In particular, it is of the structural equation modeling type. Also, the primary data collection methods in this research are a combination of library and field research methods. Furthermore, given that the study investigates relationships among all variables, correlational techniques and statistical tests will be employed to analyze the extent to which changes in one or more factors are associated with changes in others.

3.2. Population and sample

Given the limited and uncertain size of the target population, the required sample size of 384 participants was determined using the Cochran formula. The sample was subsequently selected via convenience sampling. Initially, the questionnaire, accompanied by a formal request letter, was distributed to 450 active auditors across various auditing institutions. Of the distributed instruments, 269 were returned electronically and 125 in hard copy; 10 printed questionnaires were unusable. The instrument design organized the items for each variable consecutively. To ensure data validity, several embedded items designed to detect inattentive or dishonest responses were included to screen the usability of the returned questionnaires.

3.3. Questionnaire structure

As previously noted, the questionnaire was used as the primary instrument for collecting primary data from auditors within the selected audit firms. The questionnaires had two parts and were accompanied by a letter requesting that participants complete the questionnaires and participate in the survey. The first section of the questionnaire collected demographic and general participant information, including age, gender, work experience, educational attainment, and position within the audit firm. The second section comprised specialized questions measured using an eight-item scale assessing the dependent variable, auditor impartiality, on a five-point Likert scale. Also, in the second part, the independent and moderating variables were as follows: 1 - Auditors' conscientiousness variable, including eight items. 2 - Auditors' extraversion variable, including 10 items. 3 - Power distance moderating variable, including six items.

To align the questionnaire with the research objectives, various scales were adopted from different sources. The target population consisted of auditors at all levels of audit institutions. Data collection took place in the second half of 2024. Following data collection, descriptive statistics and Harman's

single-factor test were performed using SPSS, while measurement model evaluation and hypothesis testing were conducted using SmartPLS.

At the same time, PLS-SEMs were preferred over covariance-based SEMs (CBSEMs) for three reasons. First, although this is a confirmatory study, it has a predictive lean, and PLS-SEM is more appropriate for research with a prediction-focused orientation (Hair et al., 2019). Second, the complex moderation effect of interest is simply said to be easier to estimate using PLS-SEM. In third place, PLS-SEM is not based on normality assumptions and thus is more robust than CB, which is better suited to normally distributed data. This condition is rarely the case in the social and behavioral sciences (Trail et al., 2024).

3.4. Reliability and validity

The auditor impartiality scale was measured using the audit ethics questionnaire developed by Zarefar and Zarefar (2016), which has been validated in previous studies. Furthermore, the personality scales measuring conscientiousness and extraversion originally developed by Costa and McCrae (1992) were validated in the Iranian context by Komasi et al. (2026). Similarly, the power distance items, drawn from Hofstede and Bond's (1984) cultural questionnaire, were validated and reliable, as confirmed by Beugelsdijk et al. (2017). Regarding the translation procedure, all scales were translated from English to Persian by a bilingual expert and back-translated by another independent expert to ensure linguistic equivalence.

Following prior research that established the scales' validity and reliability, we further assessed the measurement model in the current study. Table 1 presents the outer loadings, t-values, and significance levels for all indicators. The results show that most indicators have loadings above 0.50, with all t-values exceeding 1.96 ($p < 0.001$), confirming convergent validity in our sample. Also, according to the criteria of Hair et al. (1998), factor loadings of 0.30 are considered the minimum acceptable level for practical significance, provided the sample size is sufficiently large ($n \geq 350$). Since the sample size of this study is 384 respondents, the factor loadings ranging from 0.30 to 0.40 are statistically significant. Therefore, these items were retained for their theoretical importance and contribution to the instrument's construct validity.

Common method bias (CMB) was assessed using Harman's single-factor test (Table 2) (Podsakoff et al., 2003). All 32 items were entered into an unrotated principal component factor analysis. At the same time, the unrotated first factor appeared to explain only 26.54% of the variance, which is below the conventionally low commonality threshold of 50%. Additionally, eight factors with eigenvalues above 1 were identified in the analysis, but none yielded a single-factor solution that accounted for the variance structure. This result supports the conclusion that no single factor accounts for most of the covariance between items. A possible explanation for this is that common method bias would not be a major issue in the current study (Podsakoff et al., 2003; Williams et al., 1989). However, we recognize that a single-factor test is often too lenient an approach to diagnosing common method variance and that it may still be present. Future research should replicate these findings using multi-source data or longitudinal designs to further mitigate CMV concerns.

Table 1. Factor loadings of the research constructs

Variables	question	Outer Loadings	T-value	P-value
Power distance	Q207	0.628	10.956	0.000
Power distance	Q208	0.778	24.651	0.000
Power distance	Q209	0.771	22.117	0.000
Power distance	Q210	0.649	12.244	0.000
Power distance	Q211	0.656	12.171	0.000
Power distance	Q212	0.433	4.990	0.000
Extraversion	Q272	0.370	5.833	0.000
Extraversion	Q273	0.536	10.267	0.000
Extraversion	Q275	0.663	16.459	0.000
Extraversion	Q276	0.708	22.776	0.000
Extraversion	Q277	0.536	10.860	0.000
Extraversion	Q278	0.716	19.161	0.000
Extraversion	Q279	0.668	16.965	0.000
Extraversion	Q281	0.569	11.335	0.000
Extraversion	Q282	0.664	16.311	0.000
Extraversion	Q283	0.493	9.292	0.000
Conscientiousness	Q308	0.614	12.248	0.000
Conscientiousness	Q309	0.622	14.230	0.000
Conscientiousness	Q311	0.721	25.410	0.000
Conscientiousness	Q312	0.710	17.860	0.000
Conscientiousness	Q314	0.736	26.789	0.000
Conscientiousness	Q315	0.720	24.294	0.000
Conscientiousness	Q317	0.688	18.718	0.000
Conscientiousness	Q320	0.667	17.613	0.000
Auditor impartiality	Q420	0.683	20.616	0.000
Auditor impartiality	Q421	0.300	4.096	0.000
Auditor impartiality	Q422	0.730	23.367	0.000
Auditor impartiality	Q423	0.752	27.831	0.000
Auditor impartiality	Q424	0.772	23.594	0.000
Auditor impartiality	Q425	0.687	17.941	0.000
Auditor impartiality	Q426	0.775	27.938	0.000
Auditor impartiality	Q427	0.741	28.377	0.000

The reliability of the questionnaire and the study's variables was examined using Cronbach's alpha. The reliability coefficient for each section exceeded 0.7, and the overall questionnaire reliability was 0.9, confirming the data's reliability. This coefficient serves as a key criterion for assessing a scale's consistency. As is clear from Table 3, according to the sample of this study, most of the research variables have a Cronbach's alpha higher than 0.7, and the internal reliability of the constructs and the composite reliability related to the variables are also higher than 0.7, indicating that the validity and reliability of the variables are confirmed.

Table 2. Harman's single-factor test for common method bias

Component	Total	% of Variance	Cumulative %
1	8.493	26.540	26.540
2	2.535	7.922	34.462
3	1.921	6.003	40.465
4	1.611	5.035	45.500
5	1.281	4.003	49.503
6	1.183	3.698	53.201
7	1.006	3.143	56.343
8	0.925	2.892	59.235
9	0.913	2.855	62.090
10	0.876	2.738	64.828
11	0.839	2.622	67.450
12	0.776	2.426	69.876

13	0.741	2.316	72.192
14	0.706	2.207	74.398
15	0.672	2.099	76.498
16	0.638	1.992	78.490
17	0.598	1.868	80.358
18	0.593	1.852	82.210
19	0.585	1.827	84.036
20	0.546	1.705	85.742
21	0.515	1.611	87.352
22	0.463	1.445	88.798
23	0.455	1.421	90.219
24	0.443	1.385	91.604
25	0.416	1.300	92.904
26	0.385	1.204	94.108
27	0.374	1.168	95.276
28	0.366	1.144	96.421
29	0.339	1.059	97.479
30	0.290	0.905	98.384
31	0.268	0.838	99.222
32	0.249	0.778	100.000

Extraction Method: Principal Component Analysis.

Table 3. Cronbach's alpha and goodness-of-fit criteria for research variables

	Cronbach's Alpha	Internal Reliability of Constructs	Composite Reliability
Power distance	0.769	0.775	0.820
Conscientiousness	0.839	0.844	0.876
Extroversion	0.799	0.812	0.846
Auditor impartiality	0.839	0.863	0.878

4. Research findings

The research findings are presented in three main sections: demographic information, hypothesis test results, and additional peripheral findings.

4.1. Demographic data results

As detailed in Table 4, the sample comprised 384 participants, of whom 54.4% were male. Furthermore, approximately 51.8% held a bachelor's degree or higher, nearly 45% possessed two or more years of auditing experience, and over 58.6% were under 30 years of age. These descriptive statistics suggest that the sample exhibits sufficient diversity to permit generalization to the broader population.

Table 4. Demographic information of respondents

Feature	Number	Percentage	Feature	Number	Percentage
Gender			Work Experience		
Female	175	45.600	Less than 2 years	166	43.200
Male	209	54.400	Between 2 and 5 years	136	35.400
Age			Between 5 and 10 years	54	14.100
Under 30 years	225	58.600	More than 10 years	28	7.300
Between 30 and 40 years	110	28.700	Responsibility in the Audit Firm		
Between 40 and 50 years	37	9.600	Assistant Auditor	89	23.200
Over 50 years	12	3.100	Auditor	155	40.400
Education Level			Senior Auditor	67	17.500
Diploma and less than a diploma	37	9.600	Audit Supervisor	26	6.700
Associate	43	11.200	Senior Audit Supervisor	16	4.200
Bachelor	199	51.800	Audit Manager	10	2.500
Master	94	24.500	Audit Partner	21	5.500
PhD	11	2.900	***	***	***

4.2. Hypothesis testing results

The results of the hypothesis testing include descriptive statistics and inferential statistics. The descriptive statistics section is presented in two parts: Table 5 details the descriptive statistics for the research variables, and Table 6 outlines the percentage distribution of participants across the defined levels of each variable. Following this, the inferential statistics are reported sequentially, beginning with the Pearson correlation coefficients, followed by the model goodness-of-fit statistics, and concluding with the results from the structural equation modeling.

4.2.1. Descriptive statistics information

4.2.1.1. Descriptive statistics information of the variables

Table 5. Descriptive statistics of research variables

	Power distance	Conscientiousness	Extraversion	Auditor impartiality
Mean	63.870	69.530	66.310	74.450
Median	62.500	70.300	66.200	75.000
Mode	50.000	50.000	57.500	75.000
SD	17.900	16.970	15.700	17.140
Variance	320.430	288.070	246.240	293.81
Skewness	0.078	-0.104	0.043	-0.294
Kurtosis	0.192	-0.240	-0.248	-0.528
Min	0.000	0.000	15.000	6.250
Max	100	100	100	100

All variables were measured on a five-point Likert scale (1-5) and then converted to a percentage scale (0-100) for descriptive reporting. As presented in Table 5, among the variables examined, impartiality recorded the highest mean (74.45) and median (75). In contrast, power distance showed the lowest mean (63.87). Regarding variability, extroversion had the lowest standard deviation (15.70), while power distance exhibited the highest (17.90).

4.2.1.2. Results of the percentage of participants' opinions on each variable

Table 6. Percentage distribution of responses across each variable

	Very Low (%)	Low (%)	Somewhat-medium (%)	High (%)	Very high (%)	Standard deviation
Power distance	1.000	4.700	38.800	38.800	16.700	18.120
Conscientiousness	0.300	0.500	32.000	37.500	29.700	17.420
Extroversion	0.300	4.400	37.200	39.100	19.000	17.980
Auditor impartiality	0.300	0.500	25.800	34.600	38.800	18.500

Table 6 below summarizes the five categories (Very Low, Low, Somewhat-medium level, High, and Very high) for each variable, presented as percentages of the total sample of 384 respondents. For example, of the people rating themselves on conscientiousness as Somewhat-medium (about 123 respondents), 32.0% were chosen.

The response rates for conscientiousness, extraversion, and impartiality were the lowest (0.3%) at the 'Very Low' level, while power distance had the highest (1.0%). Power distance was the highest percentage (4.7%) of the joint rate within the 'Low' category, while impartiality and conscientiousness

had the lowest percentages (0.5%). The most common at the ‘Somewhat-medium’ level was power distance (38.8%), and impartiality occurred the least often (25.8%). Extraversion (39.1%) and even entertainment had a higher proportion of ‘High’ level, while impartiality only had a 34.6% share as ‘High’ level characters, respectively. Finally, in the ‘Very High’ category, impartiality recorded the highest percentage (38.8%), while power distance recorded the lowest (16.7%).

4.2.2. Inferential statistics

4.2.2.1. Pearson correlation

Table 7. Pearson correlation between research variables

	Power distance	Conscientiousness	Extraversion	Auditor impartiality
Power distance	1			
Conscientiousness	0.234**	1		
Extroversion	0.304**	0.547**	1	
Auditor impartiality	0.306**	0.567**	0.503**	1

**Significance at the 1% alpha level

Table 7 shows that, at 0.567 and 0.306, respectively, the dependent variable of impartiality has the strongest and lowest correlation with the conscientiousness and power distance variables. Additionally, the association between impartiality and conscientiousness has the highest correlation in the table (0.567), while the power distance and conscientiousness variables have the lowest correlation (0.234).

4.2.2.2. FIT Model

SRMR, or the standard root mean square residual index, is the difference between the observed correlation and the correlation matrix of the structural model. According to Table 8, if the value of this index is less than 0.08, it indicates a good fit of the Model. In this realization, the values are appropriate and accurate.

Table 8. Research Model fitness indices

Suitable size	Saturated Model
0.08<	0.068 SRMR

To assess discriminant validity, we applied the Fornell-Larcker criterion ([Fornell and Larcker, 1981](#)). This criterion requires that the square root of the average variance extracted (AVE) for each latent construct be greater than the construct's highest correlation with any other construct. Table 9 presents the Fornell-Larcker matrix, where the diagonal elements represent the AVEs and the off-diagonal elements represent inter-construct correlations. The results show that each AVE (diagonal) is larger than the corresponding off-diagonal correlations in the same row and column, thereby confirming discriminant validity.

Table 9. Fornell-Larcker matrix for research variables

	Extroversion	Auditor impartiality	Power distance	Conscientiousness
Extroversion	0.602			
Auditor Impartiality	0.532	0.696		
Power distance	0.361	0.379	0.662	
Conscientiousness	0.567	0.590	0.321	0.686

In addition to the Fornell-Larcker criterion, we assessed discriminant validity using the Heterotrait-Monotrait (HTMT) ratio of correlations (Henseler et al., 2015). Some researchers prefer the HTMT criterion because of its higher sensitivity in detecting discriminant validity issues. A common threshold for acceptable discriminant validity is HTMT < 0.85 (conservative) or HTMT < 0.90 (liberal). As shown in Table 10, all HTMT values are below 0.90, confirming that discriminant validity is established.

Table 10. Divergent validity based on the HTMT index

	Auditor impartiality	Extroversion	Power distance	Conscientiousness
Auditor impartiality				
Extroversion	0.623			
Power distance	0.408	0.402		
Conscientiousness	0.680	0.677	0.329	

Stone-Geisser's Q^2 criterion was used to assess the predictive relevance of the structural model (Stone, 1974). According to Henseler et al. (2009), Q^2 values of 0.02, 0.15, and 0.35 indicate weak, medium, and strong predictive relevance, respectively. As shown in Table 11, the Q^2 value for the endogenous construct (auditor impartiality) is 0.30, indicating that the model has medium-to-strong predictive relevance.

Table 11. Q^2 criterion for the endogeneity variable of the research

Variable	Q^2
Auditor impartiality	0.300

4.2.2.3. Structural Equation Modeling

The study's use of structural equation modeling is depicted in Figures 1 and 2. The standard regression coefficients for each association have been provided, as shown. The study's hypothesis findings are also discussed.

The adjusted coefficient of determination (R^2) is 0.448, indicating that the independent and moderating variables collectively explain 44.8% of the variance in auditor impartiality. The independent variable, auditors' conscientiousness, has a positive and significant impact on Auditors' impartiality, as shown in Figure 2, with a standard regression coefficient of 0.403 and a significance level of 0.000. This validated the first hypothesis. The second hypothesis of the study is confirmed by Figure 2, which shows that the independent variable of auditors' extraversion, with a standard regression coefficient of 0.241 and a significance level of 0.000, generally has a positive and significant effect on Auditors' impartiality. The third hypothesis of the study is rejected because, as Figure 2 illustrates, the independent variable, power distance in auditing firms, has a positive and substantial impact on Auditors' impartiality, with a standard regression coefficient of 0.190 and a significance level of 0.000.

As shown in Figure 2, the interaction effect between conscientiousness and power distance on

auditor impartiality is not statistically significant ($\beta = -0.052$, $p = 0.204$). Therefore, H4 is not supported, indicating that power distance does not moderate the relationship between conscientiousness and impartiality. Furthermore, the interaction between extraversion and power distance is statistically significant ($\beta = -0.098$, $p = 0.019$), supporting H5. This negative coefficient indicates that power distance weakens the positive relationship between extraversion and auditor impartiality.

In addition, Figure 3 presents the interaction plot for the moderating effect of power distance on the relationship between conscientiousness and auditor impartiality. As shown, although the slopes differ slightly across different levels of power distance, the interaction effect is not statistically significant ($\beta = -0.052$, $p = 0.204$). Therefore, H4 is not supported.

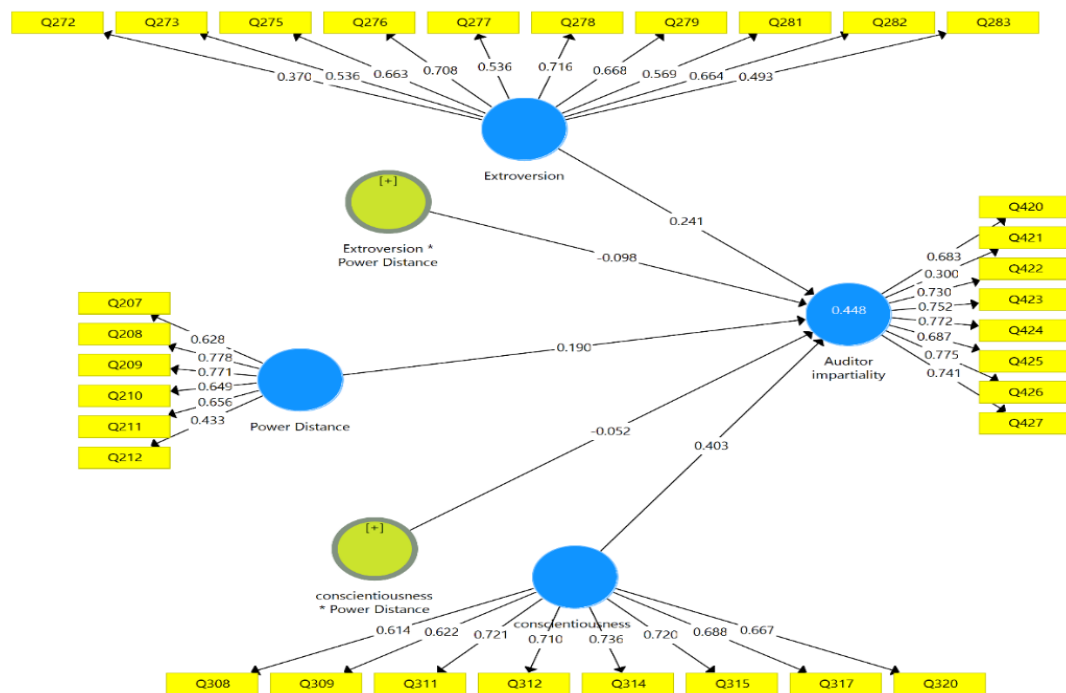


Figure 1. Standard regression coefficients of the main research Model

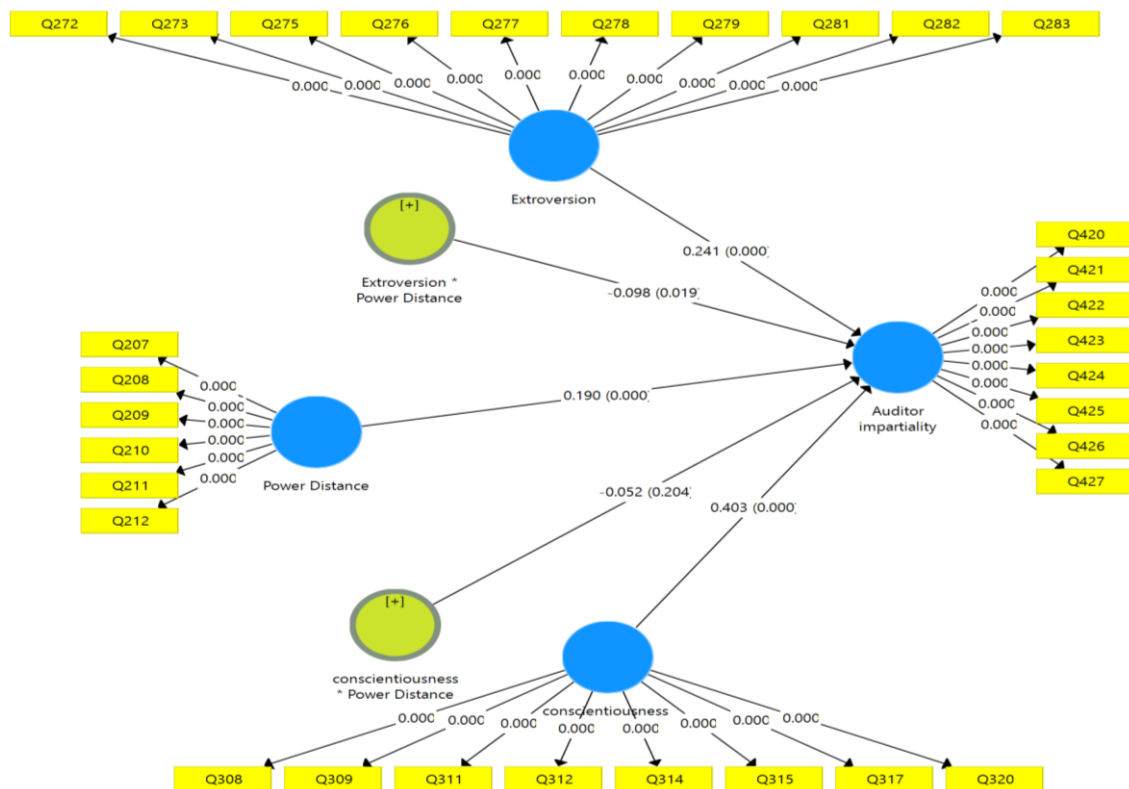


Figure 2. Significance level of standard regression coefficients of the main research Model

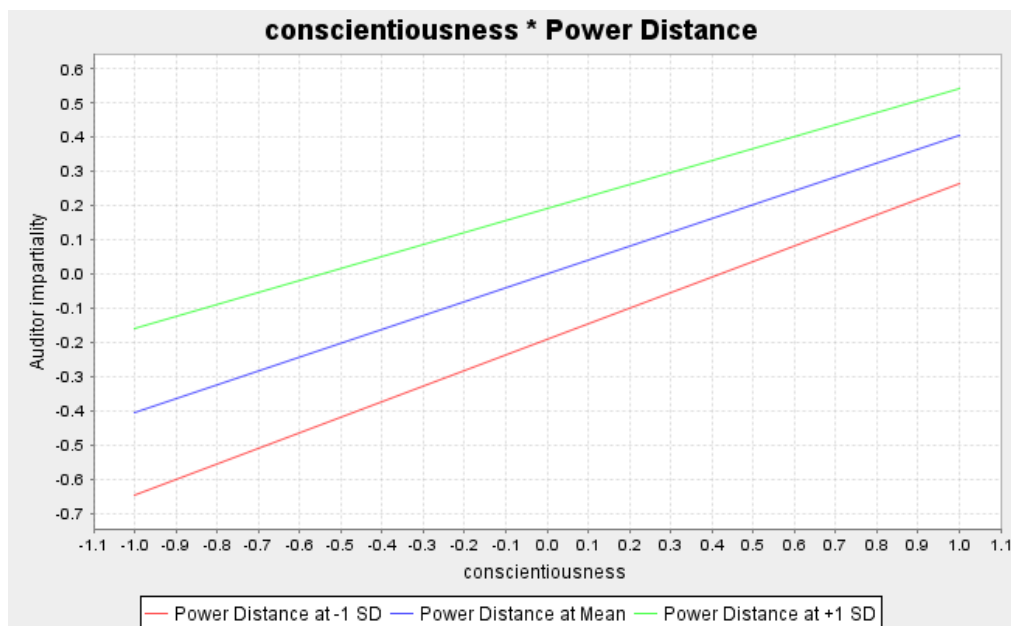


Figure 3. The moderating effect of power distance on the relationship between conscientiousness and auditor impartiality

Figure 4 presents the interaction plot for the moderating effect of power distance on the relationship between extraversion and auditor impartiality. The three lines represent different levels of power distance: low (-1 SD), mean, and high (+1 SD). As shown, the positive relationship between extraversion and impartiality is strongest when power distance is low (the top line) and weakest when power distance is high (the bottom line). This pattern confirms that power distance negatively

moderates (weakens) the extraversion-impartiality relationship, supporting H5 ($\beta = -0.098$, $p = 0.019$). All coefficients reported are standardized (β).

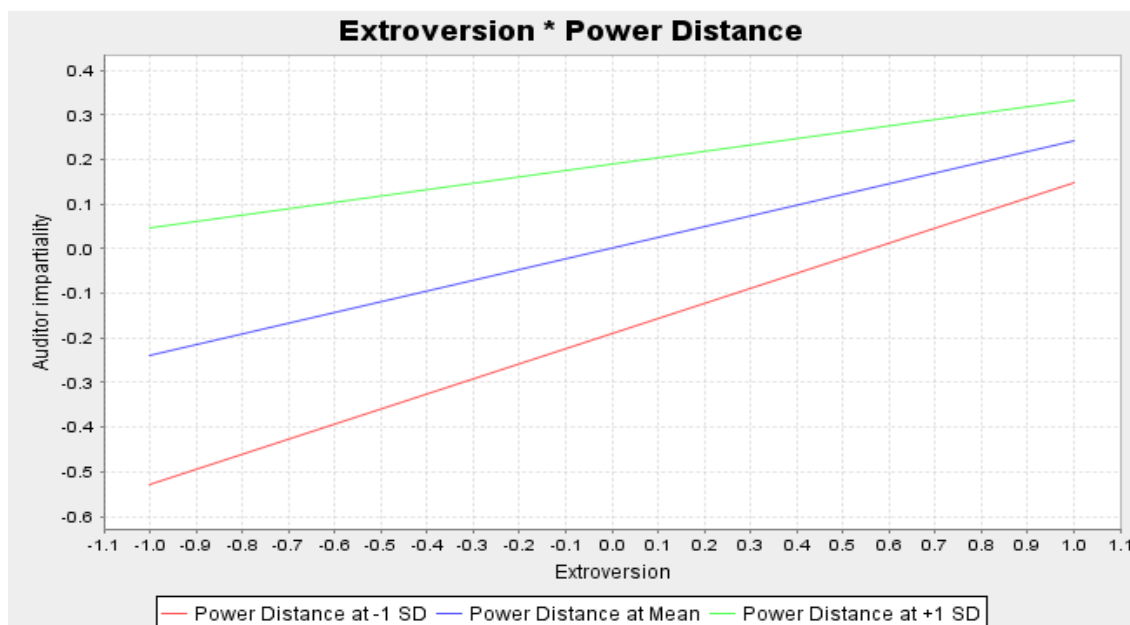


Figure 4. The moderating effect of power distance on the relationship between extraversion and auditor impartiality

4.3. Supplementary analyses

In addition, to investigate significant differences in the means of the research variables by gender, we used the analysis of variance (ANOVA) test. The results of the analysis of variance test, according to Table 12, show that since the significance of the analysis of variance test for the variables of power distance, conscientiousness, and auditor impartiality is greater than 0.05, it can be stated that there is no considerable difference between the means of the aforementioned variables for women and men in the research sample. However, there is a significant difference between the means of the auditors' extraversion variable for women and men in the research sample.

Table 12. One-way ANOVA test for the gender variable

Variable	Sources	Sum of Squares	DF	Comparison of Squares	F	Sig	Result
Power distance	Intergroup	417.820	1	417.820	1.305	0.254	Not significant, and power distance does not differ between men and women.
	Intragroup	122307.829	382	320.178			
	Total	122725.649	383				
Conscientiousness	Intergroup	2.313	1	2.313	0.008	0.929	Not significant, and conscientiousness does not differ between men and women.
	Intragroup	110329.718	382	288.821			
	Total	110332.031	383				
Extroversion	Intergroup	1324.201	1	1324.201	5.440	0.020	Significant differences in extraversion exist between men and women.
	Intragroup	92986.671	382	243.421			
	Total	94310.872	383				
Auditor impartiality	Intergroup	261.124	1	261.124	0.888	0.346	Not significant, and auditor impartiality does not differ between men and women.
	Intragroup	112271.199	382	293.904			
	Total	112532.323	383				

5. Discussion and conclusion

The main objective of this study is to examine the moderating effect of power distance on the relationship between auditors' conscientiousness and extraversion and auditors' impartiality. Overall, it demonstrates that the power distance variable moderates these correlations and confirms the relevance of the study's independent factors for auditors' impartiality.

The results of the first hypothesis show that auditors' conscientiousness has a positive and significant effect on auditor impartiality. This finding indicates that auditors who are more responsible, precise, organized, and committed to performing their duties correctly are more likely to demonstrate impartiality in their judgments and reporting. This significant relationship suggests that personality characteristics may serve as an internal mechanism supporting impartial judgment. From a practical perspective, this finding suggests that audit firms may benefit from considering conscientiousness when hiring, training, and evaluating. However, these implications should be interpreted with caution, as the study design does not permit causal inference. The findings are consistent with [Samagaio and Felício \(2022\)](#), [Wijaya and Cheisviyanny \(2024\)](#), and [Saraswati and Latrini \(2023\)](#).

The results of the second hypothesis show that auditors' extraversion has a positive and significant effect on auditor impartiality. This finding suggests that extroverted auditors, who typically possess stronger social networks, express their views more clearly, engage in more active communication, and demonstrate greater resilience to pressures in audit settings. Their ability to communicate effectively with stakeholders, ask challenging questions, and defend their professional positions may help them maintain impartiality in ambiguous situations. This finding indirectly suggests that developing communication and social skills, alongside technical knowledge, may support auditors' ability to maintain impartiality. Nevertheless, this interpretation requires further empirical testing. The findings are consistent with [Noegroho et al. \(2025\)](#) and [Chen et al. \(2023\)](#) and are indirectly supported by [Suryandari et al. \(2026\)](#).

The findings of the third hypothesis indicate a positive correlation between cultural power distance and auditor impartiality, suggesting that higher power distance is associated with greater impartiality in auditor performance. Several post hoc explanations may account for the rejection of H3. First, in high-power-distance cultures such as Iran, auditors may respond to hierarchical structures not with passive obedience but with a heightened reliance on formal rules and professional standards. This rule-based compliance could actually enhance impartiality, explaining the unexpected positive relationship. Additionally, Iranian auditors may interpret 'respect for authority' as 'respect for professional standards,' which could positively relate to impartiality. Second, the use of self-report questionnaires for all variables may have introduced common method bias. Auditors who perceive high power distance might also systematically rate themselves as more impartial due to social desirability or consistency bias, artificially inflating the observed positive correlation. Third, reverse causality is possible. Although our hypothesis assumed power distance as the independent variable, more impartial auditors might perceive their organizational hierarchy differently, possibly as more structured and rule-oriented, leading to higher reported power distance. These explanations are speculative and should be tested in future research. Also, the variable measurement tool is both an important factor and a limitation that may explain these unexpected findings.

Based on the theoretical discussions and empirical findings of this study, it can be argued that Iran has a long path ahead in developing an audit culture. A key obstacle on this path is high power distance, which suppresses upward communication, critical feedback, and professional skepticism, thereby undermining auditor impartiality. Therefore, traversing this path requires a strong cultural foundation that reduces power distance and prioritizes professional autonomy and impartiality over hierarchical conformity. The results for this hypothesis are inconsistent with both the study's

theoretical framework and the findings of [Al-Asmakh et al. \(2024\)](#) and [Shahzad et al. \(2025\)](#).

The results for H4 and H5 indicate that power distance moderates the relationship between auditors' personality characteristics (conscientiousness and extraversion) and their impartiality. However, this moderating effect is not statistically significant for conscientiousness, while it is significant for extraversion. Simple slopes analysis revealed that the positive effect of extraversion on impartiality is significant at low levels of power distance, but becomes non-significant at high levels of power distance, and higher power distance weakens the relationship between extraversion and impartiality, consistent with the negative interaction coefficient ($\beta = -0.098$, $p = 0.019$). These findings imply that organizational culture influences auditor impartiality in a differential and selective manner, contingent upon specific personality dimensions. In high power-distance cultures, hierarchical structures and norms of deference to authority can suppress the very extraverted characteristics, such as assertiveness, willingness to challenge, and comfort with risk-taking in communication, that are often conducive to impartiality. Consequently, extroverted auditors may find their capacity to maintain an independent and impartial stance toward management or senior partners diminished. To mitigate this effect, the promotion of extroverted auditors to managerial positions must be contingent upon their demonstrated ability to navigate the cultural constraints of high power distance while preserving impartiality. Furthermore, audit firms operating in such environments should develop targeted training programs for extroverted auditors. These programs ought to focus on enhancing skills in respectful negotiation with superiors, articulating dissenting views without overtly challenging authority, and establishing secure channels for reporting undue pressures.

Furthermore, it can be concluded that the core characteristics of conscientiousness (including hard work, responsibility, and commitment to standards) are more resilient even in the face of structural pressures stemming from high power distance. In other words, audit impartiality that is rooted in a deep sense of professional conscientiousness is less affected by a hierarchical culture. The characteristic of conscientiousness fosters a stronger internal locus of control in auditors. Consequently, the resulting impartiality, grounded in professional ethics and a commitment to standards, appears relatively resilient to the hierarchical pressures inherent in high-power-distance cultures. That is, conscientious auditors view impartiality not as an aggressive stance (which may be suppressed by power distance) but as a non-negotiable commitment to fully comply with auditing laws and standards. Therefore, this personality dimension is more resistant to cultural erosion. It can also be concluded that the power distance variable has a significant moderating effect on only one of the auditors' personality dimensions, namely extraversion. This underscores the necessity of a contingency perspective, positing that cultural factors do not uniformly influence all personality characteristics. Instead, their most pronounced effect is on characteristics directly associated with assertiveness and the challenging of authority, such as extraversion.

A limitation of this study is the omission of potential control variables such as age, gender, auditor rank, audit experience, and education level from the structural model. While we conducted ANOVA tests to examine gender differences (see Table 12), which indicated no significant differences for power distance, conscientiousness, and impartiality (only extraversion showed a gender difference), we acknowledge that ANOVA does not control for the simultaneous effects of these variables on the hypothesized relationships. Therefore, future research should incorporate these demographic variables as controls in the structural model (e.g., using SmartPLS) to ensure that demographic characteristics do not confound the observed effects of personality characteristics and power distance on impartiality.

References

1. Al-Asmakh, S., Elamer, A.A. and Uadiale, O. (2024) 'Cultural dynamics and tenure trajectories: How auditor tenure and culture influence key audit matters in the GCC', *Journal of Accounting Literature*, 48(3), pp. 786 - 815. <https://doi.org/10.1108/JAL-04-2024-0081>
2. Anwar, C. M. (2017). Linkages between personality and knowledge-sharing behavior in the workplace: The mediating role of affective states. *Economics and Management*, 20(2), pp. 102-115. <https://doi.org/10.15240/tul/001/2017-2-008>
3. Asnaashari, H., Safarzadeh, M. H. and Beygpanah, B. (2025). An investigation of the Iranian audit firms' transparency: evidence from an Islamic context. *Journal of Islamic Accounting and Business Research*, 16(2), pp. 249-272. <https://doi.org/10.1108/JIABR-01-2023-0003>
4. Baral, G. (2024). Delving into the happiness of professional accountants: Examining the interplay between personality traits and job and life satisfaction. *İktisadi İdari ve Siyasal Araştırmalar Dergisi (İKTİSAD)*, 9(25), pp. 727-739. <https://doi.org/10.25204/iktisad.1485201>
5. Beugelsdijk, S., Kostova, T. and Roth, K. (2017). An overview of Hofstede-inspired country-level culture research in international business since 2006. *Journal of International Business Studies*, 48(1), pp. 30-47. <https://doi.org/10.1057/s41267-016-0038-8>
6. Bik, O. and Hooghiemstra, R. (2018). Cultural differences in auditors' compliance with audit firm policy on fraud risk assessment procedures. *Auditing: A Journal of Practice & Theory*, 37(4), pp. 25-48. <https://doi.org/10.2308/ajpt-51998>
7. Bowling, N. C., Vafeiadou, A., Hammond, C., and Banissy, M. J. (2024). Extraversion and adult attachment dimensions predict attitudes towards social touch. *Journal of Research in Personality*, 111, A. 104514. <https://doi.org/10.1016/j.jrp.2024.104514>
8. Chen, Y. H., Wang, K. J. and Liu, S. H. (2023). How personality traits and professional skepticism affect auditor quality? A quantitative model. *Sustainability*, 15(2), pp. 1-15. <https://doi.org/10.3390/su15021547>
9. Costa Jr, P. T. and McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders*, 6(4), pp. 343-359. <https://doi.org/10.1521/pedi.1992.6.4.343>
10. Curtis, M. B., Conover, T. L. and Chui, L. C. (2012). A cross-cultural study of the influence of country of origin, justice, power distance, and gender on ethical decision making. *Journal of International Accounting Research*, 11(1), pp. 5-34. <https://doi.org/10.2308/jiar-10213>
11. Dust, S., Wang, P., Rode, J., Wu, Z. and Wu, X. (2021). The effect of leader and follower extraversion on leader-member exchange: An interpersonal perspective incorporating power distance orientation. *The Journal of Social Psychology*, 161(6), pp. 714-730. <https://doi.org/10.1080/00224545.2020.1848774>
12. Festinger, L. (1957). *A theory of cognitive dissonance*. Row, Peterson. Stanford, CA: Stanford University Press, Stanford
13. Fornell, C. and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), pp. 39-50. <https://doi.org/10.2307/3151312>
14. Gold, A., Hunton, J. E. and Gomaa, M. I. (2009). The impact of client and auditor gender on auditors' judgments (retracted). *Accounting Horizons*, 23(1), pp. 1-18. <https://doi.org/10.2308/acch.2009.23.1.1>
15. Haghighi, R., Bagherpour, M. A., Ghanaei Chamanabad, A. and Abbaszadeh, M. R. (2024). Are Auditors really independent in making professional judgment?. *Iranian Journal of Accounting, Auditing and Finance*, 8(3), pp. 111-130. <https://doi.org/10.22067/ijaaf.2024.44946.1432>
16. Hair, J. F., Risher, J. J., Sarstedt, M. and Ringle, C. M. (2019). When to use and how to report

- the results of PLS-SEM. *European Business Review*, 31(1), pp. 2-24. <https://doi.org/10.1108/EBR-11-2018-0203>
17. Hair, J. F., Tatham, R. L., Anderson, R. E. and Black, W. (1998). *Multivariate data analysis* (5th ed.). Prentice-Hall, New Jersey, US.
 18. Henseler, J., Ringle, C. M. and Sarstedt, M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science*, 43(1), pp. 115-135. <https://doi.org/10.1007/s11747-014-0403-8>
 19. Henseler, J., Ringle, C. M. and Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing*, 20, pp. 277–320. [https://doi.org/10.1108/S1474-7979\(2009\)0000020014](https://doi.org/10.1108/S1474-7979(2009)0000020014)
 20. Hessari, H., Daneshmandi, F., Busch, P. and Smith, S. (2025). Mitigating cyberloafing through employee adaptability: the roles of temporal leadership, teamwork attitudes and competitive work environment. *Asia-Pacific Journal of Business Administration*, 17(2), pp. 303-336. <https://doi.org/10.1108/APJBA-02-2024-0065>
 21. Hofstede, G. (1980). *Culture's Consequences: International Differences in Work-Related Values*. SAGE, University of Maastricht, Maastricht, Netherlands
 22. Hofstede, G. (1984). *Culture's consequences: International differences in work-related values* (Vol. 5). SAGE, University of Maastricht, Maastricht, Netherlands.
 23. Hofstede, G. and Bond, M. H. (1984). Hofstede's culture dimensions: An independent validation using Rokeach's value survey. *Journal of Cross-Cultural Psychology*, 15(4), pp. 417-433. <https://doi.org/10.1177/0022002184015004003>
 24. Huang, X. and Van De Vliert, E. (2003). Where intrinsic job satisfaction fails to work: National moderators of intrinsic motivation. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 24(2), pp. 159-179. <https://doi.org/10.1002/job.186>
 25. International Ethics Standards Board for Accountants (IESBA). (2018). Handbook of the international code of ethics for professional accountants. <https://www.ethicsboard.org/>
 26. Jin, X. (2024). The role of effort in understanding academic achievements: Empirical evidence from China. *European Journal of Psychology of Education*, 39(1), pp. 389-409. <https://doi.org/10.1007/s10212-023-00694-5>
 27. Kerckhofs, L., Vandenhaute, M. L. and Hardies, K. (2024). Partner narcissism in a private market setting: Consequences for audit reporting decisions and audit pricing. *International Journal of Auditing*, 28(3), pp. 500-521. <https://doi.org/10.1111/IJAU.12339>
 28. Khurana, I. K., Li, B., Yeung, K. and Yu, E. J. (2026). Audit partners' cultural trust and audit outcomes. *Journal of Accounting and Economics*, 1(1), A. 101872. <https://doi.org/10.1016/j.jacceco.2026.101872>
 29. Komasi, S. (2026). An updated validation of Persian versions of six personality questionnaires: TEMPS-A, AFFECTS, TCI, PANAS, NEO-FFI, PDQ-4. *Indian Journal of Psychological Medicine*, 48(1), pp. 49-56. <https://doi.org/10.1177/02537176241276741>
 30. Koslowsky, M., Baharav, H. and Schwarzwald, J. (2011). Management style as a mediator of the power distance-influence tactics relationship. *International Journal of Conflict Management*, 22(3), pp. 264-277. <https://doi.org/10.1108/10444061111152964>
 31. Laland, K. N. and Hoppitt, W. (2003). Do animals have culture?. *Evolutionary Anthropology: Issues, News, and Reviews: Issues, News, and Reviews*, 12(3), pp. 150-159. <https://doi.org/10.1002/evan.10111>
 32. Liegl, S. and Furtner, M. R. (2023). Introverted and yet effective? A faceted approach to the relationship between leadership and extraversion. *Frontiers in Psychology*, 14, A. 1185271.

- <https://doi.org/10.3389/fpsyg.2023.1185271>
33. Nasirpour, A., Jabbarzadeh Kangarloie, S., Bahri Sales, J. and Badavar Nahandi, Y. (2023). Audit quality model based on moral atmosphere and spirituality in Iran. *Iranian Journal of Accounting, Auditing and Finance*, 7(2), pp. 61-81. <https://doi.org/10.22067/ijaaf.2023.42476.1144>
 34. Nelson, M. W. (2009). A model and literature review of professional skepticism in auditing. *Auditing: a Journal of Practice & Theory*, 28(2), pp. 1-34. <https://doi.org/10.2308/aud.2009.28.2.1>
 35. Noegroho, Y. A. K., Tinungki, G. M., Hartono, P. G. and Matitaputty, J. S. (2025). Personality, cognitive style and auditors' professional skepticism. *Review of Behavioral Finance*, 17(5), pp. 938-959. <https://doi.org/10.1108/RBF-04-2025-0131>
 36. Nolder, C. and Riley, T. J. (2014). Effects of differences in national culture on auditors' judgments and decisions: A literature review of cross-cultural auditing studies from a judgment and decision making perspective. *Auditing: A Journal of Practice & Theory*, 33(2), pp. 141-164. <https://doi.org/10.2308/ajpt-50657>
 37. Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y. and Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), pp. 879-903. <https://doi.org/10.1037/0021-9010.88.5.879>
 38. Qian, J., Han, Z., Wang, H., Li, X. and Wang, Q. (2014). Power distance and mentor-protégé relationship quality as moderators of the relationship between informal mentoring and burnout: evidence from China. *International Journal of Mental Health Systems*, 8(1), A. 51. <https://doi.org/10.1186/1752-4458-8-51>
 39. Rhodes, P. G., Andiyasari, A. and Riantoputra, C. D. (2021). Merangkul karyawan berorientasi power distance tinggi untuk menampilkan voice behavior: Peranan managerial openness. *Jurnal Psikologi Sosial*, 19(1), pp. 39-48. <https://doi.org/10.7454/jps.2021.05>
 40. Ríos Figueroa, C. (2013). Efecto del conocimiento cultural en la percepción de los auditores: controles internos y fraude (the effect of cultural knowledge in the perception of auditors: internal controls and fraud). *Revista Internacional Administración & Finanzas*, 6(2), pp. 15-32.
 41. Ryan, R. M. and Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, 55(1), pp. 68-78. <https://doi.org/10.1037/0003-066X.55.1.68>
 42. Samagaio, A. and Felício, T. (2022). The influence of the auditor's personality in audit quality. *Journal of Business Research*, 141, pp. 794-807. <https://doi.org/10.1016/j.jbusres.2021.11.082>
 43. Saraswati, P. B. and Latrini, M. Y. (2023). Big five personality traits dan skeptisisme profesional auditor. *E-Jurnal Akuntansi*, 33(6), pp. 1628-1641. <https://doi.org/10.24843/EJA.2023.v33.i06.p16>
 44. Schwartz, S. H. (2012). An overview of the Schwartz theory of basic values. *Online Readings in Psychology and Culture*, 2(1), A. 11. <http://dx.doi.org/10.9707/2307-0919.1116>
 45. Shahzad, K., Naz, F., Iqbal, R. and Chaudhary, R. (2025). When despotic leadership is more detrimental for employee job outcomes? Role of individual-level power distance orientation. *Journal of Organizational Effectiveness: People and Performance*, 12(3), pp. 519-536. <https://doi.org/10.1108/JOEPP-12-2023-0547>
 46. Steel, P., Schmidt, J. and Shultz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin*, 134(1), pp. 138-161. <https://doi.org/10.1037/0033-2909.134.1.138>
 47. Stone, M. (1974). Cross-validators choice and assessment of statistical predictions. *Journal of the Royal Statistical Society: Series B (Methodological)*, 36(2), pp. 111-133.

- <https://doi.org/10.1111/j.2517-6161.1974.tb00994.x>
48. Suryandari, N. N. A., Bana, E., Ernawatiningsih, N. P. L. and Bhegawati, D. A. S. (2026). The Role of Auditor Personality in Detecting Corporate Fraud. *Corporate Law & Governance Review*, 8(1), A. 8. <https://doi.org/10.22495/clgrv8i1p1>
 49. Svanberg, J. and Öhman, P. (2016). Does ethical culture in audit firms support auditor objectivity?. *Accounting in Europe*, 13(1), pp. 65-79. <https://doi.org/10.1080/17449480.2016.1164324>
 50. Tarighi, H., Salehi, M., Moradi, M. and Zimon, G. (2022). Social capital, intellectual capital, and audit fee: Conflicting evidence from Iran. *Economies*, 10(2), A. 39. <https://doi.org/10.3390/economies10020039>
 51. Tiliopoulos, N., Pallier, G. and Coxon, A. P. (2010). A circle of traits: A perceptual mapping of the NEO-PI-R. *Personality and Individual Differences*, 48(1), pp. 34-39. <https://doi.org/10.1016/j.paid.2009.08.010>
 52. Trail, G. T., Kim, Y. K. and Alfaro-Barrantes, P. (2024). A critical assessment for sport management research: comparing PLS-SEM and CB-SEM techniques for moderation analysis using formative measures. *Journal of Global Sport Management*, 9(1), pp. 248-268. <https://doi.org/10.1080/24704067.2022.2098802>
 53. Veledar, B., Gadžo, A. and Jašarević, L. (2026). The influence of personal, professional, legal, and cultural factors on the application of the code of ethics in the accounting profession. *Journal of Forensic Accounting Profession*, 6(1), pp. 1–24. <https://doi.org/10.2478/jfap-2025-0006>
 54. Verma, G. G., Ganesh, K. N. and Sahishnu, M. (2025). Ethical work climate and organizational citizenship behavior in collectivist culture: does power distance matter?. In *Evidence-based HRM: a Global Forum for Empirical Scholarship* 13(3), pp. 535-552. <https://doi.org/10.1108/EBHRM-01-2024-0019>
 55. Wang, S., Zhang, L., Yang, Z., Liu, S., Wang, X., Li, Y. and Zhao, W. (2025). The unique characteristics of extraversion: A systematic review and coordinate-based meta-analysis of resting state functional magnetic resonance imaging studies. *Brain Research Bulletin*, 229, A. 111454. <https://doi.org/10.1016/j.brainresbull.2025.111454>
 56. Wijaya, L. C. and Cheisviyanny, C. (2024). Pengaruh big five personality terhadap skeptisisme profesional auditor. *Jurnal Eksplorasi Akuntansi*, 6(3), pp. 1110-1126. <https://doi.org/10.24036/jea.v6i3.1624>
 57. Williams, L. J., Cote, J. A. and Buckley, M. R. (1989). Lack of method variance in self-reported affect and perceptions at work: reality or artifact?. *Journal of Applied Psychology*, 74(3), pp. 462-468. <https://doi.org/10.1037/0021-9010.74.3.462>
 58. Winterich, K. P. and Zhang, Y. (2014). Accepting inequality deters responsibility: How power distance decreases charitable behavior. *Journal of Consumer Research*, 41(2), pp. 274-293. <https://doi.org/10.1086/675927>
 59. Zaman, R. (2024). When corporate culture matters: The case of stakeholder violations. *The British Accounting Review*, 56(1), A. 101188. <https://doi.org/10.1016/j.bar.2023.101188>
 60. Zarefar, A. and Zarefar, A. (2016). The influence of ethics, experience and competency toward the quality of auditing with professional auditor scepticism as a moderating variable. *Procedia-Social and Behavioral Sciences*, 219(114), pp. 828-832. <https://doi.org/10.1016/j.sbspro.2016.05.074>



RESEARCH ARTICLE

Board Characteristics, Audit Fees, and Political Connections: An Integration of Resource Dependence and Agency Theories

Ali Sayadi*

Department of Accounting, Faculty of Management and Economics, Lorestan University, Khorramabad, Iran

Iraj Mirhossini

Department of Accounting, Faculty of Financial Sciences, Kharazmi University, Tehran, Iran

Reza Khajezadeh, Saman Mirabbasi

Department of Accounting, Faculty of Management and Economics, Lorestan University, Khorramabad, Iran

How to cite this article:

Sayadi, A. , Mirhossini, I. , Khajezadeh, R. and Mirabbasi, S. (2026). Board Characteristics, Audit Fees, and Political Connections: An Integration of Resource Dependence and Agency Theories. *Iranian Journal of Accounting, Auditing and Finance*, 10(3), 163-185. doi: 10.22067/ijaaf.2026.48282.1650
https://ijaaf.um.ac.ir/article_48282.html

ARTICLE INFO

Abstract

Article History

Received: 2025-10-26

Accepted: 2026-02-14

Published online: 2026-07-12

Keywords:

Audit Fees, Board
Independence, Board Financial
Expertise, Political
Connections

Audit risk reduction depends on the pivotal role of effective boards. However, auditors' professional judgment is influenced by political connections, potentially impairing audit quality. The purpose of this study was to examine the impact of board characteristics on audit fees. Specifically, the moderating role of political connections will be examined. This study evaluated a sample of 107 companies listed on the Tehran Stock Exchange (TSE) from 2016 to 2023. To test the research hypotheses, pooled or panel data regression models were employed. Based on the empirical results, increasing board independence and financial expertise are associated with lower audit fees. Although political connections moderated the relationship between board independence and audit fees, they had no significant effect on the relationship between board expertise and audit fees. Political connections also directly affect the relationship between board independence and audit fees by altering corporate governance mechanisms and increasing demand for high-quality independent audits. Strict audit regulations indicated a strong institutional environment, and in such an environment, the impact of political connections on auditing is likely to be reduced. Therefore, political connections cannot have a significant moderating effect on the relationship between board expertise and audit fees. This study integrated board characteristics and political connections in an emerging-market context, thereby advancing the audit literature. Practical insights are provided for regulators, auditors, and corporate governance reformers.

 <https://doi.org/10.22067/ijaaf.2026.48282.1650>



NUMBER OF REFERENCES
78



NUMBER OF FIGURES
1



NUMBER OF TABLES
9

Homepage: <https://ijaaf.um.ac.ir>
E-ISSN: 2717-4131
P-ISSN: 2588-6142

*Corresponding Author: Ali Sayadi
Email: sayadi.a@lu.ac.ir
Tel: 09183446024
ORCID: 0009-0000-4874-0583

1. Introduction

This study examines the relationship between corporate governance and audit costs in Iran by integrating agency theory and resource dependence theory. The impact of board characteristics on audit fees for firms listed on the Tehran Stock Exchange is assessed with political connection as a moderator. The literature records wide agency costs that lead to extreme information asymmetry in many emerging markets ([Almarayeh, 2024](#)) between controlling owners and minority shareholders. Thus, the greater a firm's agency costs, the further apart its managers and owners will be ([Jensen and Meckling, 1976](#)). In jurisdictions where majority owners also hold executive positions, this problem is particularly severe because owners can pursue private benefits at the expense of minority shareholders. This leads agency theory to advocate for the use of more closely associated monitoring mechanisms to control such opportunistic behavior. It is, therefore, not surprising that corporate governance frameworks have emerged to mitigate these conflicts and improve accountability ([Daily et al., 2003](#)). [Young \(2000\)](#) demonstrated that skilled boards enhance the reliability and transparency of financial reporting and argued that effective boards reduce the risk of auditors issuing erroneous opinions. By responding to stakeholders' demand for reliable information, the board plays a central governance role in the financial reporting process. [Zaman et al. \(2011\)](#) contend that one of the board's primary functions is to preserve auditor independence and to reduce the risk of incorrect audit opinions. On the other hand, [Carcello et al. \(2002\)](#) similarly find that effective boards better fulfill oversight responsibilities and strengthen audit performance. As a result, the need for a broader audit scope to ensure audit quality may increase, leading to higher audit fees.

However, boards have often been accused of incompetence and lack of engagement, hallmarks of breaches that lead to corporate failures. The challenges have encouraged more exploration of how boards function and greater scrutiny of board composition. Estimates of the influence of political connections on firm-level governance are ambiguous, with some studies finding a negative effect. For example, [Boubakri et al. \(2013\)](#) claimed that politically connected individuals may engage in political pursuits at the expense of shareholders. Moreover, [Guedhami et al. \(2014\)](#) contended that politically connected actors could prioritize political ends over shareholders' interests. As such, political connections can drive corporate decision-making and influence outcomes such as audit quality. To improve governance and financial integrity, it is important to understand how those relationships shape board effectiveness and audit practices. The effects of political connections have been extensively examined in the literature. On one hand, a group of researchers shows that establishing political connections is generally beneficial because it provides various advantages, including direct contracts, government subsidies, tax advantages, bank loans, and regulatory benefits ([Faccio et al., 2006](#)), and easier access to finance ([Khwaja and Mian, 2005](#)). For instance, [Guedhami et al. \(2014\)](#) emphasized that managers may seek higher-quality reporting to reassure foreign investors. They find that politically connected firms hire auditors with more institutional reputation (reflecting investor demand for transparency). Conversely, political connections allow politicized actors to extract rents, which limit efficient firm performance ([Boubakri et al., 2008](#)).

Furthermore, [La Porta et al.'s \(1998\)](#) warnings in the literature include the possibility that politicians on corporate boards may divert resources and engage in earnings management, making foreign investors unwilling to invest. Recent evidence also associates board political connections with lower firm performance ([La Rocca et al., 2022](#); [Shahzad et al., 2021](#)). Since the impact of political connections on audit fees, a measure of audit effort and quality of earnings, shows mixed findings in the literature, this issue needs further investigation. According to [Chaney et al. \(2011\)](#), politically connected firms tend to exhibit lower earnings quality. This study reported that firms obtain political connections through low-earnings-quality political channels ([Chaney et al., 2011](#)). Additionally,

lower earnings quality is found for politically connected firms (Gul, 2006). Thus, political connections might increase audit risk and lead to bigger audit fees in some contexts.

On the contrary, some researchers (e.g., Cheung and Leung, 2024) have found that political connections can reduce audit fees because auditors are less likely to question politically connected clients. Perry et al. (2025) found that political pressure can discredit auditors and prevent them from reviewing information rigorously; this occurs when the political context leads to either low performance or inflated scope and fees for audit effort. Moreover, Political connections may also change the incentives for reporting from accuracy to political ends. This misalignment reduces the reliability of financial statements and diminishes governance and accountability. In addition, political connections threaten the independence of audits as auditors may avoid stricter verification to prevent consequences or the loss of clients (Guedhami et al., 2014). Accordingly, political connections moderate the relationship between board characteristics and audit fees.

In the complex and dynamic environment of Iran's capital market, determining audit fees as one of the primary mechanisms of corporate governance plays a vital role in reducing information asymmetry and conflicts of interest. However, theoretical tensions between Agency Theory (which emphasizes monitoring and control) and Resource Dependence Theory (which focuses on acquiring critical resources from the environment) create ambiguity regarding boards' behavior in auditor selection and audit fee determination. Although the accounting literature has extensively examined the determinants of audit fees, in Iran's capital market, characterized by concentrated ownership structures and the prominent presence of public and state-owned entities, traditional explanations based solely on Agency Theory (e.g., board independence and expertise) are insufficient to explain the complexities governing the pricing of audit services. In situations where firms resort to political connections to secure special privileges, reduce monitoring costs, or access scarce resources, the tension between the board's monitoring and instrumental roles intensifies.

This research is necessary to clarify whether board qualitative characteristics maintain their monitoring function in an environment heavily influenced by political pressure and integrate these interactions, or whether political motives subsume them. Answering this question is of undeniable importance to capital market policymakers and the audit profession, as it aims to enhance corporate governance standards and reduce agency costs, especially in emerging economies where political influence over economic sectors is pronounced. This study offers three key innovations compared to prior research: First, unlike many previous studies in Iran that have examined board characteristics and audit fees solely from an Agency Theory perspective, this paper adopts an integrated approach (combining Agency Theory and Resource Dependence Theory) to provide a more comprehensive explanation of board behaviors. Second, the moderating role of political connections as a contextual variable in the relationship between board characteristics and audit fees represents the main innovation of this study; specifically, it investigates whether corporate political influence neutralizes or reinforces the monitoring incentives of specialized and independent board members. Third, this research provides novel evidence to the limited body of literature on political connections at the national level from an empirical perspective, by focusing on the institutional context of Iran's capital market, given its specific corporate governance features and ownership structure, which can significantly contribute to a better understanding of the interactions between corporate governance and political structures in developing countries.

Thus, this study examines how board independence and board financial expertise relate to audit fees and whether political connections moderate these relationships. Specifically, it investigates the following two research questions:

Research Question 1: What is the nature of the relationship between board independence and audit fees?

Research Question 2: What is the nature of the relationship between board expertise and audit fees?

Research Question 3: How do political connections influence the relationship between board independence and audit fees?

Research Question 4: How do political connections influence the relationship between board expertise and audit fees?

2. Literature review and hypothesis development

The academic literature has paid particular attention to the relationship between corporate governance and audit quality, particularly regarding the implications of board structure on audit fees. Audit fees are often considered a measure of audit quality, with higher fees generally indicating a more comprehensive audit process. DeAngelo (1981) defines audit quality as the likelihood that an auditor may detect and report a violation in the client's accounting system. Similarly, Palmrose (1988) defines audit quality as the level of assurance provided by the auditor to users of financial statements. Almarayeh et al. (2020) suggested that effective corporate governance structures are designed to increase accountability and transparency, ultimately affecting the quality of financial reporting and the audit process. Previous research has highlighted two dominant views on the relationship between corporate governance and audit fees. On the one hand, stronger corporate governance may lead to higher audit fees, as firms may seek broader audit services to mitigate risks associated with weak corporate governance. On the other hand, firms with stronger corporate governance mechanisms may experience lower audit fees because the reduced likelihood of financial misstatements reduces auditors' risk. Numerous studies support the first argument, claiming that effective corporate governance requires a reputable audit firm, additional audit tasks, and, consequently, better audit quality, which in turn leads to higher audit fees (Al-Hajaya, 2019; Lin, 2018).

2.1. Audit fees

Audit fees are the amounts paid by the client to the audit firm under the contract for the provision of audit services (Carcello et al., 2002). This variable is important from the perspectives of management and shareholders (for the reliability of financial reporting) and audit firms (for revenue maximization and maintaining competitiveness). Empirical evidence indicates that corporate governance mechanisms, including ownership structure and managerial independence, influence audit fees through their effects on agency problems, risk, firm size, and complexity. Nevertheless, the relative importance of these factors has been shown to vary across countries (Ali and Lesage, 2013). The regulatory context and the demand structure for audit services differ markedly in developing countries, leading to direct implications for how audit fees are set (Kamarudin et al., 2022).

2.2. Board independence

Board independence is the absence of material relationships or interests between a board member and the company, which constitutes one dimension of the theoretical aspect of corporate governance (Durisin and Puzone, 2009). In particular, if anyone has an employment relationship with the firm or serves as its executive director, they are not independent (Kosnik, 1987). Since the board of directors is an essential piece of the corporate governance system, profiling its structure and functioning is particularly key (Weisbach and Hermalin, 2003). Board independence is widely seen as one of the cornerstones of effective corporate governance, helping board members fulfill their legal responsibilities (monitoring company management) more effectively and better protect shareholder interests (Bozec and Dia, 2017). In this context, the non-executive directors function more effectively in their supervisory role because of their independent position and objective view towards

management decisions and actions (Kosnik, 1987). While financial market participants and regulators alike focus strongly on board independence, empirical evidence is somewhat mixed (Weisbach and Hermalin, 2003). However, there are still debates amongst regulators and corporate governance code providers, with some pointing out that board independence is still often viewed as a proxy for board effectiveness.

2.3. Board financial expertise

Board expertise refers to the knowledge and experience of the management team, which can have either positive or negative effects on firm performance. This expertise, particularly in the financial domain, is often measured using criteria such as work experience, professional experience, and academic qualifications. Given that the stock exchange market requires financial knowledge and expertise to make investment decisions, a manager with a sufficient understanding of financial literature, investment strategies, and the implications of financing strategies is less likely to exhibit irrational behavior. On the contrary, a manager with a finance background may increase cash-flow investment sensitivity (Gupta et al. 2021). The data linking board members' skills and a firm's operating performance support this intuition, with greater educational attainment and functional knowledge leading to better project selection, better execution, and ultimately better operating performance (Khan et al., 2025). Even though some explanations from the human relations school, like resource dependence theory, argue that what really accounts for firm achievement is not the individual ability of managers nor their knowledge or skills but factors such as product quality and life cycle (Hannan and Freeman, 1977), Human capital theory outlines that managerial knowledge and ability can be important antecedents of organizational performance (Coff, 2002). The board of directors, being the highest body in the firm, is responsible for developing and promoting high-level strategies, making major decisions, and managing overall company operations. The assumption is that members who lack sufficient accounting or financial knowledge are less sensitive to the truthfulness of reports, and that the presence of a member with adequate financial experience makes other members more sensitive and alert (Ghazalat et al., 2017). The results of numerous empirical studies conducted in other countries also indicated that good corporate governance, including board financial expertise, leads to better firm performance (Balatbat et al., 2004).

2.4. Political Connections

Political connections of economic enterprises, considered one of the factors influencing the quality of financial reporting, can be defined at the firm level as relationships between board members and individuals close to centers of power and government (Xiao et al., 2023). According to political economy theory, economics and politics have a reciprocal influence on one another, such that political decisions directly and indirectly affect economic activities (Pabst and Scazzieri, 2023). Evidence suggests that in developing countries, political connections are even more important than a firm's economic fundamentals in determining profitability, because firms with political connections, through easier access to credit and financial resources from state-owned banks, are less dependent on reported earnings for financing (Boubakri et al., 2012). The indicators of such political connections include the presence of board members affiliated with the government or political institutions, a government or semi-government major shareholder, and a higher long-term debt ratio resulting from access to low-cost, long-term resources (Lee and Wang, 2017).

2.5. Independence of the board and audit fees

In accounting research, the relationship between board independence and audit fees is of great importance. To analyze this issue, both the demand-side and supply-side perspectives on audit services should be considered. From a demand perspective, increased board independence aimed at more effective monitoring requires exerting greater pressure to achieve audit objectives. This demand for greater assurance by the board is logical only when the resulting benefits outweigh the audit costs (Carcello et al., 2002). Board independence stems from its members' independence, which is essential for ensuring board effectiveness by maintaining their independence from management. Current non-executive directors focus more on ensuring audit quality and reducing threats to auditor independence. Furthermore, independent board members, due to their higher motivation to preserve professional reputation and reduce legal liability risk, demand higher-quality audits to ensure the accuracy of financial statements. Board independence enhances audit quality and, as a signal of a strong internal control environment, encourages auditors to assess control risks more carefully (Krishnan and Visvanathan, 2009). Beasley (1996) argued that greater oversight by a highly independent board reduces the occurrence of financial statement fraud. Many scholars have highlighted scrutiny regarding independent members' disclosure on boards (Al Farooque et al., 2020; Brickley and James, 1987). Board member independence, a significant monitoring mechanism, is illustrated by Carcello and Neal (2003). Since independent directors are not dependent on management, they claim that these independent directors have more ability to question management when necessary. Uang et al. (2006) argue that independent directors may reduce agency costs, narrow the gap between related parties and independent auditors, and effectively improve reporting quality. In addition, Abbott and Parker (2000) showed that firms characterized by a higher proportion of independent directors use high-quality auditors. Moreover, Carcello et al. (2002) found that non-executive board members are more likely to invest in higher quality to "protect their reputational capital, minimize litigation exposure, and promote shareholders' interests." Similarly, Jizi and Nehme (2018) found that boards with more independence determined their actual audit fees. Kalia et al. (2023) showed that board characteristics (size and independence) are positively related to audit fees. Since outside directors are more motivated to enhance the quality of financial reporting, boards with more non-executive directors may signal greater demand for higher-quality audit services, thereby positively impacting audit fees. The first hypothesis, ensuing from these findings, is the following:

H1: Board independence has a positive effect on audit fees.

2.6. Board of directors' expertise and audit fees

Board members with expertise in finance play a vital role in improving corporate performance. Carcello et al. (2002) emphasized that a board with expert members provides more effective oversight. DeFond et al (2005) also stated that managers who pay attention to financial matters can better monitor the reporting process. Various studies (Bédard et al., 2004) have shown that board members' expertise and experience are directly related to board effectiveness. Carcello et al. (2002) reported a positive correlation between financial expertise and audit fees. Saggese et al. (2023) also found that board human capital policies and regulations can affect audit quality and financial reporting. Consequently, the expertise of board members in auditing and finance increases audit quality and its fees. Based on this evidence, the second hypothesis of the study is formulated as follows:

H2: Board expertise has a positive effect on audit fees.

2.7. The relationship between Political connection, audit fees, and board independence, based on Resource Dependence and Agency Theories

To explicate this subject, two theoretical frameworks, Agency Theory and Resource Dependence Theory, have been employed in an integrated manner. Agency Theory, which is predicated on the separation of ownership from control ([Jensen and Meckling, 1976](#)), posits that managers may prioritize their personal interests over those of shareholders. Consequently, firms should strengthen their corporate governance mechanisms. In this regard, board independence is an essential monitoring mechanism that helps minimize managerial opportunism ([Haniffa and Cooke, 2002](#); [Buniamin et al., 2011](#)). Agency Theory suggests that audit fees reflect the costs of monitoring managers when shareholders are exposed to risk after entrusting management with decision-making ([Ashbaugh and Warfield, 2003](#)).

In contrast, according to Resource Dependence Theory ([Pfeffer and Salancik, 1978](#)), the board of directors acts as a mediator between the firm and external resources and, ultimately, better governance reduces dependence on the macro environment by improving firm performance. According to this theory, the firm is viewed as a bundle of resources and capabilities ([Barney, 1991](#)). Political connections are considered a form of distinct capital that reduces uncertainty and facilitates access to government benefits, making them an effective resource for the firm in this context ([Pfeffer and Salancik, 1978](#)).

The integration of these two theories in the formulation of the third hypothesis proceeds as follows: political connections serve as a nexus between the firm's external environment and the effectiveness of internal control mechanisms. In Iran's economic context, characterized by an extensive government role, concentrated ownership structures, institutional uncertainties, frequent regulatory changes, and financing constraints, political connections can play a determinative role in shaping corporate circumstances. In such an environment, board independence alone may not exert a direct and uniform impact on audit fees; rather, this effect is contingent upon the firm's network of political relationships. More precisely, when assessing risk and the quality of internal monitoring, the auditor evaluates board independence in light of the firm's political connections. Agency Theory further cautions that political connections can escalate the financial burden of political activities and diminish profitability by engendering information asymmetry and hidden costs ([Pang and Wang, 2021](#)). Drawing upon the foregoing discussion, the third hypothesis is articulated as follows:

H3: The Political connections moderate the relationship between audit fees and board independence.

2.8. The relationship between Political connection, audit fee, and board expertise, based on Resource Dependence and Agency Theories

Two theoretical approaches, Agency Theory and Resource Dependence Theory, have been employed to support this hypothesis. As a theoretical approach, Agency Theory is based on the recognition of an ongoing separation between ownership and control ([Jensen and Meckling 1976](#)), which holds that managers have interests other than those of shareholders. Therefore, it is important to improve the corporate governance mechanisms to align interests ([Almarayeh, 2024](#)). In this context, board members' knowledge is recognized as one of the foundational determinants of monitoring. [Baker et al. \(2020\)](#) highlighted that boards with high levels of technical and managerial competence exercise stronger oversight of managerial behavior, thereby increasing accountability within the firm. Furthermore, [Khatib et al. \(2021\)](#) conducted an extensive study confirming that board diversity includes board members educated in non-Western developed nations (NWDS), bringing renewed perspectives and insights to boards. According to Agency Theory, audit fees are a proxy of

monitoring expenses, and independent auditors help organizations overcome asymmetric information problems (Ashbaugh and Warfield, 2003).

In parallel with Agency Theory, Resource Dependence Theory (Pfeffer and Salancik, 1978) emphasizes that firms need mechanisms to minimize environmental uncertainty and gain access to external resources. From this perspective, the firm is conceptualized as a collection of resources and capabilities (Barney, 1991). Political connections serve as strategic resources, functioning as intangible assets for accessing government benefits and support, thereby producing a sustained competitive advantage. Ben-Amar et al. (2017) claim that board composition is an invaluable resource for creating value and helps ensure sustainable competitive advantage by increasing access to knowledgeable external resources and strategic guidance.

The combination of these two theories within the framework of the fourth hypothesis implies that board expertise, as an internal control mechanism, cannot independently and uniformly affect audit fees; rather, this effect is moderated by political connections. In other words, political connections act as a bridge between the external environment and the effectiveness of internal mechanisms. This issue is of particular importance in the Iranian economy, which is characterized by pervasive government presence, concentrated ownership, institutional instability, frequent legal changes, and financial constraints. In such a context, when assessing risk and operational complexity, the auditor evaluates board expertise not in isolation, but in interaction with the firm's political relationships. Agency Theory also cautions that political connections may impose hidden costs and influence firm performance (Shahzad et al., 2021). Therefore, it is expected that the effectiveness of board expertise in determining audit fees is contingent on the firm's level of political connections, which moderates the relationship between board expertise and audit fees. Drawing upon the foregoing discussion, the fourth hypothesis is articulated as follows:

H4: Political connections moderate the relationship between audit fees and board expertise.

3. Research methodology

To empirically test the research hypotheses, linear regression models were estimated using EViews 9 and Stata 17. The initial population comprised all firms listed on the Tehran Stock Exchange (TSE) from 2016 to 2023. A systematic screening process was applied to ensure sample homogeneity and data reliability. Specifically, firms were required to have been listed on the TSE before the start of the sample period, maintain a fiscal year ending on March 20 to ensure intertemporal and cross-sectional comparability, and not operate in financial intermediary sectors (including insurance, banking, brokerage, and investment firms), as these entities are subject to distinct regulatory and reporting frameworks. Furthermore, all necessary annual financial data for computing the study variables had to be available, and firms that changed their fiscal year during the sample period were excluded to preserve temporal consistency. The final sample, satisfying all selection criteria, consisted of 107 firm-year observations from firms listed on the Tehran Stock Exchange.

3.1. Hypothesis testing models

This study seeks to examine the relationship between board characteristics and audit fees and assess the moderating role of board political connections. The hypotheses were tested following the methodology of Almarayeh (2024). The following four models were developed to test the research

hypotheses.

Model (1) for examining hypothesis one: Board independence has a significant effect on audit fees.

$$\begin{aligned} \ln AF_{i,t} = & \alpha_0 + \alpha_1 \text{Bind}_{i,t} + \alpha_2 \text{AudCh}_{i,t} + \alpha_3 \text{AudSize}_{i,t} + \alpha_4 \text{Current}_{i,t} + \alpha_5 \text{Inv}_{i,t} + \alpha_6 \text{Lev}_{i,t} \\ & + \alpha_7 \text{Loss}_{i,t} + \alpha_8 \text{Roa}_{i,t} + \alpha_9 \text{Size}_{i,t} + \alpha_{10} \text{Year fixed effect}_{i,t} + \epsilon_{i,t} \end{aligned}$$

Model (2) for examining hypothesis two: Board expertise has a significant impact on audit compensation.

$$\begin{aligned} \ln AF_{i,t} = & \alpha_0 + \alpha_1 \text{Bexp}_{i,t} + \alpha_2 \text{AudCh}_{i,t} + \alpha_3 \text{AudSize}_{i,t} + \alpha_4 \text{Current}_{i,t} + \alpha_5 \text{Inv}_{i,t} + \alpha_6 \text{Lev}_{i,t} \\ & + \alpha_7 \text{Loss}_{i,t} + \alpha_8 \text{Roa}_{i,t} + \alpha_9 \text{Size}_{i,t} + \alpha_{10} \text{Year fixed effect}_{i,t} + \epsilon_{i,t} \end{aligned}$$

Model (3) for examining hypothesis three: The Political connections moderate the relationship between audit fees and board independence.

$$\begin{aligned} \ln AF_{i,t} = & \alpha_0 + \alpha_1 \text{Bind}_{i,t} + \alpha_2 \text{Polcon}_{i,t} + \alpha_3 \text{Bind} * \text{Polcon}_{i,t} + \alpha_4 \text{AudCh}_{i,t} + \alpha_5 \text{AudSize}_{i,t} \\ & + \alpha_6 \text{Current}_{i,t} + \alpha_7 \text{Inv}_{i,t} + \alpha_8 \text{Lev}_{i,t} + \alpha_9 \text{Loss}_{i,t} + \alpha_{10} \text{Roa}_{i,t} + \alpha_{11} \text{Size}_{i,t} \\ & + \alpha_{12} \text{Year fixed effect}_{i,t} + \epsilon_{i,t} \end{aligned}$$

Model (4) for examining hypothesis four: The Political connections moderate the relationship between audit fees and board expertise.

$$\begin{aligned} \ln AF_{i,t} = & \alpha_0 + \alpha_1 \text{Bexp}_{i,t} + \alpha_2 \text{Polcon}_{i,t} + \alpha_3 \text{Bexp} * \text{Polcon}_{i,t} + \alpha_4 \text{AudCh}_{i,t} + \alpha_5 \text{AudSize}_{i,t} \\ & + \alpha_6 \text{Current}_{i,t} + \alpha_7 \text{Inv}_{i,t} + \alpha_8 \text{Lev}_{i,t} + \alpha_9 \text{Loss}_{i,t} + \alpha_{10} \text{Roa}_{i,t} + \alpha_{11} \text{Size}_{i,t} \\ & + \alpha_{12} \text{Year fixed effect}_{i,t} + \epsilon_{i,t} \end{aligned}$$

3.2. Research variables measurement

3.2.1. Dependent variable

Audit fee (LnAf): In this study, the dependent variable is the natural logarithm of the audit fee (Hoitash et al., 2007). Information regarding the audit fee was extracted from the notes accompanying the financial statements of the administrative and general expenses section and the specified other expenses.

3.2.2. Independent variables

A) *Board independence (Bind)*: equal to the ratio of non-executive directors to the total number of board members of the company (Sanan et al., 2021).

B) *Board expertise (Bexp)*: is calculated by dividing the number of members with financial experience (Specifically, accounting, financial management, economics, and Auditing) and expertise by the total number of board members. Financial expertise refers to academic education in accounting or related fields and work experience in finance (Jeanjean and Stolowy, 2009).

3.2.3. Moderating variable

Board political connection (Polcon): Political connections are the independent variable of the present study. The government's influence and ownership in companies have been used in many previous studies (Sari and Anugerah, 2011; Dinç, 2005) to measure political connections. According

to the study by [Nguyen et al. \(2022\)](#), there is no complete disclosure of political work experience (government positions) of board members and CEOs in the Iranian study environment. In 2022, the minimum ownership percentage was set at 50 percent as a guideline for establishing political connections, having substantial influence, and implementing the business unit's financial policies. As a result, if the percentage of ownership by government, quasi-governmental, and government-affiliated entities in the company is at least 50 percent, political connections are assigned a value of 1; otherwise, 0 ([Uddin, 2016](#)).

3.2.4. Control variables

A set of control variables has been incorporated into the research model to ensure that the observed relationship between independent variables and the dependent variable is adequately explained and the effects of other factors influencing audit fees are eliminated. Larger firms typically have more complex operations, a higher volume of transactions, and a more extensive organizational structure. These complexities increase auditors' workload; consequently, firm size (Size) affects audit fees. A high level of debt in a firm's capital structure increases its financial risk. An increase in financial risk can lead to higher audit risk, as the likelihood of bankruptcy or liquidity problems intensifies. To cover this additional risk, auditors demand higher fees; thus, financial leverage (Lev) also influences audit fees. Return on assets serves as a measure of firm profitability. High profitability may indicate that the audit is beneficial for the firm (due to value added) or, conversely, may signal potential risks in income recognition (e.g., earnings management) that require greater auditor attention. Therefore, return on assets (ROA) affects audit fees. Loss-making firms often face greater challenges regarding going concern and earnings quality. These circumstances can increase audit risk, as auditors must exercise more care in examining the going concern assumption and the reasons for losses. Hence, the loss variable (Loss) can be a determinant of audit fees. When a firm changes its auditor, the new auditor may encounter challenges such as incomplete access to the previous auditor's records or the need for a deeper understanding of the firm's operations and accounts. These factors can lead to an increased workload and, consequently, higher fees in the first period of engagement with the new auditor. The current ratio is a measure of a firm's short-term liquidity, indicating the firm's ability to meet its short-term financial obligations using current resources. A high current ratio reflects a favorable liquidity position and the firm's capacity to pay short-term debts. This favorable condition reduces the risk of bankruptcy or immediate financial distress. Since auditors are responsible for assessing a firm's financial risks, a reduction in risk due to strong liquidity can decrease the workload required to evaluate those risks, thereby reducing audit fees. On the other hand, a current ratio that is either very high or very low may indicate complexities in managing current assets and liabilities. In both cases, the complexities associated with this ratio can affect the auditor's workload and fees. The current ratio (Current) in the control model helps isolate the effect of the firm's short-term liquidity position on auditors' fee-setting decisions, thereby providing a more precise analysis of the relationship between board independence and audit fees. Control variables used in this study include:

Firm size (Size), equal to the natural logarithm of assets;

Current ratio (Current), equal to current assets divided by current liabilities;

Financial leverage (Lev) is equal to total debt divided by total assets.

Return on assets (ROA) is the ratio of net income to assets, and

Inventory intensity (Inv) is the ratio of total inventory of materials and goods to total assets.

Year fixed effects are year control variables.

Three dummy variables are also used: the *loss variable (Loss)*, which takes the value 1 if the company incurs a loss and 0 otherwise; the *audit firm size variable (AudSize)*, which takes the value 1 if the auditing organization performs the audit and 0 otherwise. and the *Auditor Change variable*

(*AudCh*), which takes the value of one if the company changes its auditor in the year under review, and zero otherwise.

4. Findings

4.1. Descriptive statistics

Table 1 presents the descriptive statistics for the key variables under study.

Table 1. Descriptive statistics of quantitative variables

Variable title	Mean	Median	max	min	Std. Dev	Obs
Audit fee	7.715	7.668	10.805	5.634	0.839	854
Independence of the Board of Directors	0.648	0.600	1.000	0.000	0.186	854
Board of Directors' Expertise	0.264	0.263	0.600	0.000	0.126	854
Current ratio	1.831	1.456	27.095	0.161	1.688	854
Inventory intensity ratio	0.233	0.228	0.560	0.032	0.116	854
Financial leverage	0.578	0.538	14.588	0.027	0.722	854
Return on Assets	0.155	0.136	0.523	-0.100	0.132	854
Firm size	15.237	15.035	21.362	11.406	1.543	854

Table 2. Descriptive statistics of the dummy variables

Variable title	Mean	Median	max	min	Std. Dev	Obs
Political Connection	0.516	1.000	1.000	0.000	0.500	854
Auditor Change	0.250	0.000	1.000	0.000	0.433	854
Audit Firm Size	0.148	0.000	1.000	0.000	0.356	854
Loss	0.103	0.000	1.000	0.000	0.304	854

Tables 1 and 2 present descriptive statistics of quantitative and dummy variables of the study, including mean, median, maximum, minimum, standard deviation, and number of observations (854 firm-year observations). According to the findings, the mean Audit Fee is 7.715, close to its median (7.668), indicating a relatively symmetric distribution. Board Independence has a mean of 0.648, indicating that more than ~60% of board members are independent on average (a value higher than 0.5 indicates that only independents remain). For Board Expertise, the mean is 0.264, meaning that about one in every four board members has relevant expertise. Dummy variables (Political Connection (mean = 0.516), Auditor Change (mean = 0.250), Audit Firm Size (mean = 0.148), and Loss (mean = 0.103) show proportions of these events taking place in the sample set Inventory Intensity Ratio (mean = 0.233) and Return on Assets a balanced environment for firms or at least imply such an environment, thus high-degree of firm concentration is avoided. At the same time, Financial Leverage denotes firms' average debt, with a mean of 0.578 and a maximum of 14.588, indicating that existing companies have almost no equity, thereby creating an excessive burden on all stakeholders. Firm Size has a mean of 15.237 and a standard deviation of 1.543, indicating decent sample differences. According to [De Vaus \(2002\)](#), there is a borderline chance that the data may be normal when the total number of observations exceeds 100 (a general, rough guideline for normality).

4.2. Correlation coefficients of research variables

To determine the type and strength of the relationships between the research variables, including the dependent variable, independent variables, moderator, and control variables, Pearson correlation coefficients were calculated. The results of these calculations are presented in Table 3. The Pearson correlation coefficient ranges from -1 to +1, where positive values indicate a direct relationship and negative values indicate an inverse relationship between the variables.

Table 3. Pearson correlation coefficients of research variables

Variables	LnAf	Bind	Bexp	Polcon	AudCh	AudSize	Current	Inv	Lev	Loss	Roa	Size
LnAf	1.000											
Bind	-0.191	1.000										
Bexp	0.079	-0.052	1.000									
Polcon	0.109	-0.177	0.114	1.000								
AudCh	-0.171	0.025	-0.033	-0.078	1.000							
AudSize	0.298	-0.229	0.005	0.167	-0.234	1.000						
Current	-0.010	0.139	-0.047	-0.194	0.035	-0.124	1.000					
Inv	0.024	-0.084	0.079	0.109	0.040	-0.095	-0.036	1.000				
Lev	0.039	-0.141	0.032	0.068	-0.053	0.137	-0.205	0.050	1.000			
Loss	0.008	-0.151	0.007	-0.018	-0.000	0.128	-0.171	-0.096	0.256	1.000		
Roa	0.068	0.082	0.056	-0.015	0.006	-0.084	0.288	0.053	-0.179	-0.314	1.000	
Size	0.556	-0.174	0.126	0.057	-0.177	0.353	-0.080	-0.106	0.014	0.040	0.119	1.000

Table 3 presents the Pearson correlation matrix among the research variables. These coefficients measure the direction and strength of the linear relationship between each pair of variables. The results show that the correlation coefficient between Firm Size and Audit Fee is the highest (0.556), indicating a relatively strong positive correlation. Furthermore, Audit Firm Size has a positive and very significant connection (0.298) with audit fees. In respect of other variables, Board Independence has a negative correlation (-0.191) with audit fees, and Auditor Change likewise has a negative correlation (-0.171) with audit fees. The other correlation coefficients among the explanatory variables are modest, typically below 0.30. Importantly, all correlation coefficients between variables are below 0.70 (or 0.7). This indicates no serious multicollinearity among the model's predictor variables. Therefore, all variables can be used simultaneously in the regression analysis without concerns about multicollinearity.

4.3. Reliability test

Before estimating the model parameters, the reliability of the variables was examined using the Levine, Lin, and Chu test. The quantitative results of the reliability test of the research variables are presented in Table 4.

Table 4. Levine, Lin, and Chu Test

Description	Variable Indicator	Levine, Lin, and Chu	Sig. level
Audit fee	LnAf	-23.691	0.000
Independence of the Board of Directors	Bind	-9.948	0.000
Board of Directors' Expertise	Bexp	-6.211	0.000
Political Connection	Polcon	-2.058	0.019
Auditor Change	AudCh	-19.483	0.000
Audit Firm Size	AudSize	-16.496	0.000
Current ratio	Current	-24.317	0.000
Inventory intensity ratio	Inv	-20.354	0.000
Financial leverage	Lev	-23.826	0.000
Loss	Loss	-3.515	0.000
Return on Assets	Roa	-18.170	0.000
Firm size	Size	-18.590	0.000

As shown, the research variables are reliable at the 95% confidence level. The parameters can be estimated without worrying about being incorrect.

4.4. Endogeneity Test

The relationship between the independence of the Board of Directors, the Board of Directors' expertise, and audit fees may be endogenous. The endogeneity test was used to inform robustness checks with instrumental variables (IV) to address concerns about reverse causality. In other words, the accounting comparability variable in regression models should be exogenous. Generally, in the econometric literature, an endogenous variable is strongly correlated (collinear) with the remaining components of the model. Therefore, in this study, the Durbin-Wu-Hausman test was used to examine the exogeneity and the absence of reverse causality. In this method, the model is first fitted using two-stage least squares (TSLS). Then, the exogeneity of the accounting comparability variable is examined by using the control and independent variables as instruments. The null hypothesis in this test is that the variable is exogenous, and the alternative hypothesis is that the variable is endogenous. If the significance level obtained for a variable is greater than 5%, there is no evidence of endogeneity, and the null hypothesis of exogeneity is not rejected. The results of this test are presented in Table 5. The significance level obtained for the comparability variable is less than 5%, so the accounting comparability variable is endogenous.

Table 5. Endogeneity test

Variables	Model number	Significance level
Bind	Model (1)	0.033
Bind	Model (2)	0.042
Bexp	Model (3)	0.000
Bexp	Model (4)	0.002

4.5. Hypothesis Test

Given that the data used in this research are panel (firm-year) and can be either pooled or panel, the Limmer F-test is used to select between the pooled and panel data methods for model estimation. To examine the Limpert F-test, if the probability of the F-statistic is greater than 0.05, the pooled data method should be used. Otherwise, the panel data method should be selected. If the panel data method is accepted, then the Hausman test should be conducted to choose between the random effects and fixed effects models. The Hausman test assesses whether the estimated regression error is related to the model's independent variables. If such a relationship exists, the fixed-effects model will be used; if it does not, the random-effects model will be used. In the Hausman test, if the p-value of the Chi-square statistic is greater than 0.05, the random-effects model should be used. Otherwise, the fixed effects model should be chosen. The results of the Limer F-test and the Hausman test are presented in Table 6.

Table 6. Results of the F-limer and Hausman test

Model	Limer F-test			Hausman test		
	Test statistic	Probability	Result	Test statistic	Probability	Result
First	6.565	0.000	a table	38.213	0.000	fixed effects
Second	6.757	0.000	a table	41.405	0.000	fixed effects
Third	6.544	0.000	a table	41.032	0.000	fixed effects
Fourth	7.380	0.000	a table	29.783	0.001	fixed effects

As observed, the significance level of the Limer F-test for the models is less than 0.05. Therefore, the panel data method has been accepted. Additionally, based on the significance level of the Hausman test, which is also less than 0.05, the fixed-effects model is accepted.

4.6. The existence of variance heterogeneity Test

To perform the test for heterogeneity of variance, the Modified Wald test was used. The results are presented in Table 7. The Modified Wald test is used to test for heteroscedasticity of variance in linear regression models and examines the dependence of the variance of the residual terms on the values of the explanatory variables. In examining heteroscedasticity of variance, statistical tests and the null and alternative hypotheses are defined as follows:

H_0 : The variance of the errors does not depend on the values of the independent variables.

H_1 : The variance of the errors does depend on the values of the independent variables.

This means that the null hypothesis states that the data are homoscedastic, and the alternative hypothesis seeks to confirm the presence of heteroscedasticity. Now, if the software output Prob > 0.05, we can reject the null hypothesis and conclude that there is no variance heteroscedasticity in the data; conversely, if Prob < 0.05, we can reject the null hypothesis and conclude that there is variance heteroscedasticity in the data.

Table 7. Modified Wald test results

Model number	chi2	Significance level
Model (1)	1.221	0.273
Model (2)	1.432	0.241
Model (3)	1.256	0.269
Model (4)	1.413	0.235

The results of the Modified Wald test indicate the absence of variance heterogeneity in the research models.

4.7. Results of hypothesis tests

In the present study, the Durbin-Watson test was used to assess autocorrelation in the employed model; if the Durbin-Watson statistic lies between 1.5 and 2.5, the independence of the model error term can be accepted. In addition, the Wooldridge test was used to detect the autocorrelation of the model. Moreover, the Fisher test was used to examine the overall significance of the employed model. Additionally, the VIFs are all less than 10, indicating no multicollinearity (Table 8).

According to Model (1) in the table above, the F-statistic is 12.636, with a significance level of 0.000, indicating that the overall regression model is statistically significant. The adjusted R-squared value is 0.644, indicating that approximately 64% of the variation in the dependent variable is explained by the independent and control variables. Furthermore, the VIF statistic is less than 10 for all variables, indicating no multicollinearity. Given that the Wooldridge statistic is greater than 5%, the assumption of independence of the model's error terms can be accepted. The value of the model's Durbin-Watson statistic is within the specified range, so the assumption of no autocorrelation in the model's error term is acceptable. The t-statistic for the first hypothesis that board independence has a significant effect on audit fees is -1.974. Therefore, the first hypothesis was not rejected at the 5% error level.

According to Model (2) in the table above, the F-statistic is 12.730 with a significance level of 0.000, indicating that the overall regression model is statistically significant. The adjusted R-squared value is

0.645, indicating that approximately 64% of the variation in the dependent variable is explained by the independent and control variables. Furthermore, the VIF statistic is less than 10 for all variables, indicating no multicollinearity. Given that the Wooldridge statistic is greater than 5%, the assumption of independence of the model's error terms can be accepted. The value of the model's Durbin-Watson statistic is within the specified range, so the assumption of no autocorrelation in the model's error term is acceptable. The t-statistic for the second hypothesis, that board expertise has a significant effect on audit fees, is -2.057. Therefore, the second hypothesis was not rejected at the 5% error level.

Table 8. Regression results for hypotheses 1 and 2

Variable	Model 1				Model 2			
	Coefficient	T -Value	Sig.	Vif	Coefficient	T -Value	Sig.	Vif
Bind	-0.309	-2.392	0.017	1.122	*	*	*	*
Bexp	*	*	*	*	-0.437	-2.057	0.040	1.408
AudCh	-0.073	-1.894	0.058	1.067	-0.069	-1.817	0.069	1.106
AudSize	0.221	0.906	0.365	1.047	0.219	0.899	0.369	1.043
Current	0.034	1.866	0.062	1.948	0.033	1.869	0.062	1.939
Inv	0.424	1.645	0.100	1.347	0.448	1.745	0.081	1.300
Lev	0.046	1.354	0.176	2.330	0.055	2.004	0.045	2.157
Loss	0.226	2.647	0.008	1.308	0.209	2.493	0.012	1.307
Roa	0.354	1.768	0.077	1.218	0.370	1.853	0.064	1.349
Size	0.399	12.749	0.000	1.225	0.403	12.976	0.000	1.372
C(constant)	1.512	2.959	0.003	NA	1.404	2.814	0.005	NA
Year fixed effect	Controlled				Controlled			
R-squared	0.699				0.700			
Adjusted R-squared	0.644				0.645			
Durbin -Watson	2.032				2.024			
Wooldridge statistic	1.503				1.634			
Prob (Wooldridge statistic)	0.239				0.231			
F-statistic	12.636				12.730			
prob (F-statistic)	0.000				0.000			

Table 9. Regression results for hypotheses 3 and 4

Variable	Model 3				Model 4			
	Coefficient	T -Value	Sig.	Vif	Coefficient	T -Value	Sig.	Vif
Bind	-0.498	-2.344	0.0194	1.907	*	*	*	*
Bexp	*	*	*	*	-0.663	-2.571	0.010	3.140
Polcon	-0.486	-2.260	0.0241	6.469	-0.188	-1.160	0.246	2.409
Bind*Polcon	0.599	2.023	0.0435	6.780	*	*	*	*
Bexp*Polcon	*	*	*	*	0.241	0.786	0.431	3.355
AudCh	-0.071	-1.845	0.0655	1.099	-0.058	-1.318	0.188	1.418
AudSize	0.243	0.998	0.3184	1.059	0.270	0.964	0.335	1.069
Current	0.035	1.933	0.0537	1.990	0.034	1.871	0.062	1.950
Inv	0.398	1.547	0.1223	1.489	0.715	2.535	0.011	1.683
Lev	0.047	1.382	0.1674	2.531	0.038	1.115	0.265	2.640
Loss	0.227	2.652	0.0082	1.421	0.133	1.100	0.271	1.611
Roa	0.360	1.797	0.0727	1.345	0.330	1.543	0.123	1.713
Size	0.395	12.533	0.0000	1.350	0.415	12.944	0.000	1.491
C(constant)	1.796	3.382	0.0008	NA	1.229	2.330	0.020	NA
Year fixed effect	Controlled				Controlled			
R-squared	0.702				0.735			
Adjusted R-squared	0.646				0.667			
Durbin -Watson	2.037				2.174			
Wooldridge statistic	1.691				1.768			
Prob (Wooldridge statistic)	0.204				0.187			
F-statistic	12.528				10.693			
prob (F-statistic)	0.000				0.000			

According to the findings in Model (3) in Table 9, the F-statistic is 12.528, with a significance level of 0.000, indicating that the overall regression model is statistically significant. The adjusted R-squared value is 0.646, indicating that approximately 64% of the variation in the dependent variable is explained by the independent and control variables. Furthermore, the VIF statistic is less than 10 for all variables, indicating no multicollinearity. Given that the Wooldridge statistic is greater than 5%, the independence of the model's error term can be accepted. The value of the model's Durbin-Watson statistic is within the specified range, so the assumption of no autocorrelation in the model's error term is acceptable. The t-statistic for hypothesis three, which states that political connections have a significant effect on the relationship between board independence and audit fees, is 2.023. Therefore, hypothesis three is not rejected at the 5% error level.

Also, referring to Model (4) in Table 9, the F-statistic is 10.693, with a significance level of 0.000, indicating that the overall regression model is statistically significant. The T-statistic for the fourth hypothesis, which states that political connections have a significant effect on the relationship between board expertise and audit fees, is 0.786. Therefore, hypothesis four is rejected at the 5% error level.

5. Discussion and conclusion

From a supply-side perspective, auditors believe that increasing board independence signals a high-control environment and conditions, leading to a lower assessment of control risk based on the results of testing the first hypothesis and in line with [Bozec and Dia \(2017\)](#) research. This reduces the use of audit procedures and audit fees. Therefore, a negative relationship is predicted between board independence and audit fees. Contrary to the research of [Abbott and Parker \(2000\)](#) and drawing on the agency theory perspective, it can be argued that an independent board, by reducing agency costs, provides greater assurance to shareholders and other stakeholders that their interests are prioritized. This assurance reduces the inherent and control risks that the auditor assesses during the audit process. According to the audit services pricing model ([Simunic, 1980](#)), audit fees are a direct function of the level of audit risk. Thus, the reduction in risk resulting from high board independence logically leads to lower audit fees. Furthermore, according to research by [Carcello et al. \(2002\)](#) and [Abbott et al. \(2004\)](#), from a substitution perspective, the presence of non-executive directors (increased board independence) has not reduced the need for auditing, due to the board's rigorous oversight.

According to the research of [Kalelkar and Khan \(2016\)](#), considering the results of testing the second hypothesis, some accounting researchers, like [Hay et al. \(2006\)](#), [Simon and Francis \(1988\)](#), and [Simunic \(1980\)](#), believe that audit fees are a function of business risk and auditor risk. Empirical evidence also indicates that auditors consider business risk and audit risk when assessing management characteristics. According to a study by [Kalelkar and Khan \(2016\)](#) examining the relationship between managers' financial knowledge and audit pricing in the U.S. capital market, auditors demand lower fees when dealing with managers with financial expertise. This research shows that when auditors become aware that company managers have financial knowledge and experience, they assess audit risk more favorably and, as a result, charge lower fees to their clients. The mechanism of this relationship is as follows: the presence of managers with financial expertise reduces inaccurate financial reporting and, consequently, lowers audit risk, thereby decreasing audit fees. In other words, the findings of this study indicate a negative relationship between managers' financial knowledge and audit fees. Moreover, consistent with our research, [Kalelkar and Khan \(2016\)](#) ultimately concluded that companies whose managers have financial knowledge pay lower audit fees.

Regarding the third hypothesis, it has been stated that given the state structure of the Iranian economy and the prominent role of the government as a shareholder in many companies, state ownership is considered an indicator of the presence of political connections on the board of directors

These political connections lead to greater control over the activities of companies with government shareholders and government agencies, and auditors, aware of this, plan a broader scope of work to protect the interests of shareholders. Consequently, audit risk increases, and auditors demand higher fees (Choi et al., 2024). On the other hand, the demand-based view of audit services indicates a significant relationship between board independence and audit fees (Hay et al., 2008). According to this view, the combination of the board's professional and legal responsibilities and the company's ownership structure, together with the need to interact directly with the internal environment, affects audit fees (O'Sullivan, 2000; Carcello et al., 2002; Fan and Wong, 2005). However, this relationship is affected in companies where the board's political connections are prominent. In other words, the presence of political connections has a direct and significant effect on the relationship between board independence and audit fees, because these connections affect corporate governance mechanisms and increase the demand for independent auditing. Therefore, the political connections of the board, as a moderating variable, affect the relationship between board independence and audit fees, indicating the key role of government ownership structure and political connections of the board in increasing the demand for higher quality audits and, consequently, increasing audit fees.

According to the results of the fourth hypothesis and based on legitimacy theory, companies can use independent, high-quality auditing methods to maintain their social legitimacy, regardless of the board's political connections. Furthermore, the effect of board expertise on the monitoring structure and quality of corporate governance is fundamental and internal to the organization. It is not necessarily influenced by external factors such as political connections. Under such circumstances, board expertise alone determines audit fees, and political connections do not play a moderating role. Moreover, the auditing profession operates based on the principles of independence and professional standards. According to Defond and Zhang (2014), audit fees are not a function of political considerations but primarily of inherent business risk, operational complexity, and the quality of internal controls. Even in the presence of board members with political connections, if the audit is conducted in accordance with professional standards, political connections may have little effect on the relationship between board expertise and audit fees. The findings of this study showed that the fourth hypothesis, which proposed a moderating role for political connections in the relationship between board expertise and audit fees, was rejected in the sample studied.

Nevertheless, the correlation between board independence and audit fees is considerably moderated by political relationships. The two board features in question differ in nature and in how they exert their impact. For successful risk mitigation and monitoring, board independence, defined as the capacity to make decisions independently of management, is crucial (Fama and Jensen, 1983; Beasley, 1996). A board's independence can be severely compromised in Iran's institutional setting due to the prevalence of political connections and the significant government presence. Board members' technical knowledge and abilities, particularly in financial and accounting matters, enable a more precise evaluation of reporting risks, more effective qualitative supervision of management performance, and a better grasp of operational complexity (Simunic, 1980). Political connections, while they may influence some operational or strategic decisions, do not alter the essence of the board's technical knowledge and specialized skills.

To improve auditors' efficiency, it is suggested that their fees be adjusted based on their level of expertise and financial knowledge relative to board members. Companies are also recommended to employ experts, experienced individuals with appropriate financial knowledge, and improve the performance of their boards. Given the board's key role in the company and the financial reporting process, its financial knowledge shapes its attitude towards auditing. A board with strong financial knowledge is likely to demand greater rigor in the audit process. Therefore, audit firms should consider board members' financial knowledge when planning and performing audits. Future research should examine how board political connections affect the relationship between product market competition and audit fees, as well as the relationship between economic policy uncertainty and audit fees.

References

1. Abbott, L. J. and Parker, S. (2000). Auditor selection and audit committee characteristics. *Auditing: A Journal of Practice & Theory*, 19(2), pp. 47-66. <https://doi.org/10.2308/aud.2000.19.2.47>
2. Abbott, L. J., Parker, S. and Peters, G. F. (2004). Audit committee characteristics and restatements. *Auditing: A journal of practice & theory*, 23(1), pp. 69-87. <https://doi.org/10.2308/aud.2004.23.1.69>
3. Al Farooque, O., Buachoom, W. and Sun, L. (2020). Board, audit committee, ownership and financial performance—emerging trends from Thailand. *Pacific Accounting Review*, 32(1), pp. 54-81. <https://doi.org/10.1108/PAR-10-2018-0079>
4. Al-Hajaya, K. (2019). The impact of audit committee effectiveness on audit quality: Evidence from the Middle East. *International Review of Management and Marketing*, 9(5), pp. 1-8. <https://doi.org/10.32479/irmm.8341>
5. Ali, C. B. and Lesage, C. (2013). Audit pricing and nature of controlling shareholders: Evidence from France. *China Journal of Accounting Research*, 6(1), pp. 21-34. <https://doi.org/10.1016/j.cjar.2012.08.002>
6. Almarayeh, T. (2024). Audit committees' independence and earnings management in developing countries: evidence from MENA countries. *Journal of Financial Reporting and Accounting*. ahead-of-print. <https://doi.org/10.1108/JFRA-11-2023-0652>
7. Almarayeh, T. S., Aibar-Guzmán, B. and Abdullatif, M. (2020). Does audit quality influence earnings management in emerging markets? Evidence from Jordan. *Revista de Contabilidad*, 23(1), pp. 64-74.
8. Ashbaugh, H. and Warfield, T. D. (2003). Audits as a corporate governance mechanism: Evidence from the German market. *Journal of International Accounting Research*, 2(1), pp. 1-21. <https://doi.org/10.2308/jiar.2003.2.1.1>
9. Baker, H. K., Pandey, N., Kumar, S. and Haldar, A. (2020). A bibliometric analysis of board diversity: Current status, development, and future research directions. *Journal of Business Research*, 108, pp. 232-246. <https://doi.org/10.1016/j.jbusres.2019.11.025>
10. Balatbat, M. C., Taylor, S. L. and Walter, T. S. (2004). Corporate governance, insider ownership and operating performance of Australian initial public offerings. *Accounting & Finance*, 44(3),

- pp. 299-328. <https://doi.org/10.1111/j.1467-629x.2004.00114.x>
11. Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), pp. 99-120. <https://doi.org/10.1177/014920639101700108>
 12. Be'dard, J., Chtourou, S. M. and Courteau, L. (2004). The effect of audit committee expertise, independence, and activity on aggressive earnings management. *Auditing: A Journal of Practice & Theory*, 23(2), pp. 13-35. <https://doi.org/10.2308/aud.2004.23.2.13>
 13. Beasley, M. S. (1996). An empirical analysis of the relation between the board of director composition and financial statement fraud. *Accounting Review*, 71(4), pp. 443-465. <https://ssrn.com/abstract=2647>
 14. Ben-Amar, W., Chang, M. and McIlkenny, P. (2017). Board gender diversity and corporate response to sustainability initiatives: Evidence from the carbon disclosure project. *Journal of Business Ethics*, 142(2), pp. 369-383. <https://doi.org/10.1007/s10551-015-2759-1>
 15. Boubakri, N., Cosset, J. C. and Saffar, W. (2008). Political connections of newly privatized firms. *Journal of Corporate Finance*, 14(5), pp. 654-673. <https://doi.org/10.1016/j.jcorpfin.2008.08.003>
 16. Boubakri, N., El Ghouli, S. and Saffar, W. (2013). Cash holdings of politically connected firms. *Journal of Multinational Financial Management*, 23(4), pp. 338-355. <https://doi.org/10.1016/j.mulfin.2013.06.002>
 17. Boubakri, N., Guedhami, O., Mishra, D. and Saffar, W. (2012). Political connections and the cost of equity capital. *Journal of Corporate Finance*, 18(3), pp. 541-559. <https://doi.org/10.1016/j.jcorpfin.2012.02.005>
 18. Bozec, R. and Dia, M. (2017). Monitoring function of the board and audit fees: contingent upon ownership concentration. *International Journal of Accounting & Information Management*, 25(1), pp. 70-90. <https://doi.org/10.1108/IJAIM-05-2016-0054>
 19. Brickley, J. A. and James, C. M. (1987). The takeover market, corporate board composition, and ownership structure: The case of banking. *The Journal of Law and Economics*, 30(1), pp. 161-180. <https://www.jstor.org/stable/725396>
 20. Buniamin, S., Alrazi, B., Johari, N. H. and Rahman, N. R. A. (2011). Corporate governance practices and environmental reporting of companies in Malaysia: Finding possibilities of double thumbs up. *Jurnal Pengurusan*, 32(2011), pp. 55-71. <https://www.researchgate.net/publication/274458534>
 21. Carcello, J. V. and Neal, T. L. (2003). Audit committee characteristics and auditor dismissals following "new" going-concern reports. *The Accounting Review*, 78(1), pp. 95-117. <https://doi.org/10.2308/accr.2003.78.1.95>
 22. Carcello, J. V., Hermanson, D. R., Neal, T. L. and Riley Jr, R. A. (2002). Board characteristics and audit fees. *Contemporary Accounting Research*, 19(3), pp. 365-384. <https://doi.org/10.1506/CHWK-GMQ0-MLKE-K03V>
 23. Chaney, P. K., Faccio, M. and Parsley, D. (2011). The quality of accounting information in politically connected firms. *Journal of Accounting and Economics*, 51(1-2), pp. 58-76. <https://doi.org/10.1016/j.jacceco.2010.07.003>
 24. Cheung, P. Y. and Leung, T. Y. (2024). Do political connections help or harm family firms? An audit pricing perspective. *The Chinese Economy*, 57(5), pp. 395-414. <https://doi.org/10.1080/10971475.2024.2321706>
 25. Choi, H., Karim, K. and Zhang, Y. (2024). Earnings management risk and audit pricing: Evidence from big bath accounting. *Journal of Business Finance & Accounting*, 51(7-8), pp. 1789-1827. <https://doi.org/10.1111/jbfa.12764>
 26. Coff, R. W. (2002). Human capital, shared expertise, and the likelihood of impasse in corporate

- acquisitions. *Journal of Management*, 28(1), pp. 107-128. <https://doi.org/10.1177/014920630202800107>
27. Daily, C. M., Dalton, D. R. and Cannella Jr, A. A. (2003). Corporate governance: Decades of dialogue and data. *Academy of Management Review*, 28(3), pp. 371-382. <https://doi.org/10.2307/30040727>
 28. De Angelo, L. E. (1981). Auditor size and audit quality. *Journal of Accounting and Economics*, 3(3), pp. 183-199. [https://doi.org/10.1016/0165-4101\(81\)90002-1](https://doi.org/10.1016/0165-4101(81)90002-1)
 29. De Vaus, D. A. (2002). *Surveys in social research* (5th ed.). London: Routledge, London, U.K. <https://parsmodir.com/wp-content/uploads/2020/05/devaus.pdf>
 30. DeFond, M. and Zhang, J. (2014). A review of archival auditing research. *Journal of accounting and economics*, 58(2-3), pp. 275-326. <https://doi.org/10.1016/j.jacceco.2014.09.002>
 31. DeFond, M. L., Hann, R. N. and Hu, X. (2005). Does the market value financial expertise on audit committees of boards of directors? *Journal of Accounting Research*, 43(2), pp. 153-193. <https://doi.org/10.1111/j.1475-679x.2005.00166.x>
 32. Dinç, I. S. (2005). Politicians and banks: Political influences on government-owned banks in emerging markets. *Journal of Financial Economics*, 77(2), pp. 453-479. <https://doi.org/10.1016/j.jfineco.2004.06.011>
 33. Durisin, B. and Puzone, F. (2009). Maturation of corporate governance research, 1993–2007: An assessment. *Corporate Governance: An International Review*, 17(3), pp. 266-291. <https://doi.org/10.1111/j.1467-8683.2009.00739.x>
 34. Faccio, M., Masulis, R. W. and McConnell, J. J. (2006). Political connections and corporate bailouts. *The Journal of Finance*, 61(6), pp. 2597-2635. <https://doi.org/10.1111/j.1540-6261.2006.01000.x>
 35. Fama, E. F. and Jensen, M. C. (1983). Separation of ownership and control. *The Journal of Law and Economics*, 26(2), pp. 301-325. <https://www.jstor.org/stable/725104>
 36. Fan, J. P. and Wong, T. J. (2005). Do external auditors perform a corporate governance role in emerging markets? Evidence from East Asia. *Journal of Accounting Research*, 43(1), pp. 35-72. <https://doi.org/10.1111/j.1475-679x.2004.00162.x>
 37. Ghazalat, A. N. M., Islam, Md. A., Noor, I. B. M. and Abu Haija, A. A. (2017). Does the competency attributes of board members' enhance the monitoring of earnings management? Evidence from a developing country. *Corporate Board: Role, Duties and Composition*, 13(1-1), pp. 100-108. <http://dx.doi.org/10.22495/cbv13i1clart4>
 38. Guedhami, O., Pittman, J. A. and Saffar, W. (2014). Auditor choice in politically connected firms. *Journal of Accounting Research*, 52(1), pp. 107-162. <https://doi.org/10.1111/1475-679X.12032>
 39. Gul, F. A. (2006). Auditors' response to political connections and cronyism in Malaysia. *Journal of Accounting Research*, 44(5), pp. 931-963. <https://doi.org/10.1111/j.1475-679X.2006.00220.x>
 40. Gupta, G., Mahakud, J. and Verma, V. (2021). CEO's education and investment–cash flow sensitivity: An empirical investigation. *Review of Behavioral Finance*, 13(5), pp. 555-577. <https://doi.org/10.1108/IJMF-01-2020-0020>
 41. Haniffa, R. M. and Cooke, T. E. (2002). Culture, corporate governance and disclosure in Malaysian corporations. *Abacus*, 38(3), pp. 317-349. <https://doi.org/10.1111/1467-6281.00112>
 42. Hannan, M. T., and Freeman, J. (1977). The population ecology of organizations. *American Journal of Sociology*, 82(5), pp. 929-964. <https://doi.org/10.1086/226424>
 43. Hay, D. C., Knechel, W. R. and Wong, N. (2006). Audit fees: A meta-analysis of the effect of supply and demand attributes. *Contemporary Accounting Research*, 23(1), pp. 141-191. <https://doi.org/10.1506/4XR4-KT5V-E8CN-91GX>

44. Hay, D., Knechel, W. R., and Ling, H. (2008). Evidence on the impact of internal control and corporate governance on audit fees. *International Journal of Auditing*, 12(1), pp. 9-24. <https://doi.org/10.1111/j.1099-1123.2008.00367.x>
45. Hoitash, R., Markelevich, A. and Barragato, C. A. (2007). Auditor fees and audit quality. *Managerial Auditing Journal*, 22(8), pp. 761-786. <https://doi.org/10.1108/02686900710819634>
46. Jeanjean, T. and Stolowy, H. (2009). Determinants of board members' financial expertise—Empirical evidence from France. *The International Journal of Accounting*, 44(4), pp. 378-402. <https://doi.org/10.1016/j.intacc.2009.09.002>
47. Jensen, M. C., and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4), pp. 305–360.
48. Jizi, M. and Nehme, R. (2018). Board monitoring and audit fees: the moderating role of CEO/chair dual roles. *Managerial Auditing Journal*, 33(2), pp. 217-243. <https://doi.org/10.1108/MAJ-10-2016-1464>
49. Kalelkar, R. and Khan, S. (2016). CEO financial background and audit pricing. *Accounting Horizons*, 30(3), pp. 325-339. <https://doi.org/10.2308/acch-51442>
50. Kalia, D., Basu, D. and Kundu, S. (2023). Board characteristics and demand for audit quality: a meta-analysis. *Asian Review of Accounting*, 31(1), pp. 153-175. <https://doi.org/10.1108/ARA-05-2022-0121>
51. Kamarudin, K. A., Wan Ismail, W. A. and Jahan, M. A. (2022). The effect of mandatory audit firm rotation and the auditing and reporting standards on the auditor competition and audit fees relationship: Evidence from emerging markets. *International Journal of Auditing*, 26(3), pp. 415-438. <https://doi.org/10.1111/ijau.12276>
52. Khan, I., Saha, A. K., Kamal, Y. and Islam, M. S. (2025). Do ownership structure and board composition matter in firm performance? Regulatory influence in developing economy. *International Journal of Disclosure and Governance*, 22(1), pp. 263-283. <https://doi.org/10.1057/s41310-024-00257-0>
53. Khatib, S. F., Abdullah, D. F., Elamer, A. A. and Abueid, R. (2021). Nudging toward diversity in the boardroom: A systematic literature review of board diversity of financial institutions. *Business Strategy and the Environment*, 30(2), pp. 985-1002. <https://doi.org/10.1002/bse.2665>
54. Khwaja, A. I. and Mian, A. (2005). Do lenders favor politically connected firms? Rent provision in an emerging financial market. *The Quarterly Journal of Economics*, 120(4), pp. 1371-1411. <https://doi.org/10.1162/003355305775097524>
55. Kosnik, R. D. (1987). Greenmail: A study of board performance in corporate governance. *Administrative Science Quarterly*, 32(2), pp. 163-185. <https://doi.org/10.2307/2393124>
56. Krishnan, G. V. and Visvanathan, G. (2009). Does the SOX definition of an accounting expert matter? The association between audit committee directors' accounting expertise and the level of audit fees. *Contemporary Accounting Research*, 26(4), pp. 1081- 1115. <https://doi.org/10.2139/ssrn.866884>
57. La Porta, R., Lopez-de-Silanes, F., Shleifer, A. and Vishny, R. W. (1998). Law and finance. *Journal of Political Economy*, 106(6), pp. 1113-1155. <https://doi.org/10.1086/250042>
58. La Rocca, M., Fasano, F., Cappa, F. and Neha, N. (2022). The relationship between political connections and firm performance: An empirical analysis in Europe. *Finance Research Letters*, 49(1), A. 103157. <https://doi.org/10.1016/j.frl.2022.103157>
59. Lee, W. and Wang, L. (2017). Do political connections affect stock price crash risk? Firm-level evidence from China. *Review of Quantitative Finance and Accounting*, 48(3), pp. 643-676. <https://doi.org/10.1007/s11156-016-0563-3>
60. Lin, Y. C. (2018). The consequences of audit committee quality. *Managerial Auditing Journal*,

- 33(2), pp. 192-216. <https://doi.org/10.1108/MAJ-03-2016-1350>
61. Nguyen, B., Do, H. and Le, C. (2022). How much state ownership do hybrid firms need for better performance?. *Small Business Economics*, 59(3), pp. 845-871. <https://doi.org/10.1007/s11187-021-00556-8>
 62. O'sullivan, N. (2000). The impact of board composition and ownership on audit quality: Evidence from large UK companies. *The British Accounting Review*, 32(4), pp. 397-414. <https://doi.org/10.1006/bare.2000.0139>
 63. Pabst, A. and Scazzieri, R. (2023). *The Constitution of Political Economy: Polity, Society and the Commonweal*. Cambridge: Cambridge University Press, Cambridge, England. <https://doi.org/10.1017/9781108923231>
 64. Palmrose, Z. V. (1988). 1987 Competitive manuscript co-winner: an analysis of auditor litigation and audit service quality. *Accounting Review*, 63(1), pp. 55-73.
 65. Pang, C. and Wang, Y. (2021). Political connections, legal environments and firm performance around the world. *International Journal of Finance & Economics*, 26(3), pp. 4393-4409. <https://doi.org/10.1002/ijfe.2021>
 66. Perry, Y. K. Z., Adrian, C. and Truong, C. (2025). Politically connected audit committee and audit risk: Evidence from a quasi-experiment. *Accounting & Finance*, 65(2), pp. 1428-1467. <https://doi.org/10.1111/acfi.13373>
 67. Pfeffer, J. and Salancik, G. R. (1978). *The external control of organizations: A resource dependence perspective*. New York: Harper & Row, New York.
 68. Saggese, S., Sarto, F., Romano, R. and Viganò, R. (2023). Implications of directors' education for audit fees: does the audit committee matter? *Managerial Auditing Journal*, 38(6), pp. 901-926. <https://doi.org/10.1108/MAJ-05-2022-3544>
 69. Sanan, N. K., Jaisinghani, D. and Yadav, S. (2021). Corporate governance, firm performance, and business group affiliation: evidence from India. *Management Decision*, 59(8), pp. 1863-1876. <https://doi.org/10.1108/MD-03-2019-0376>
 70. Sari, R. N. and Anugerah, R. (2011). The effect of political influence and corporate transparency on firm performance: empirical evidence from Indonesian listed companies. *Journal of Modern Accounting and Auditing*, 7(8), pp. 773-783. <http://www.davidpublisher.org/index.php/Home/Article/index?id=2441>
 71. Shahzad, F., Saeed, A., Asim, G. A., Qureshi, F., Rehman, I. U. and Qureshi, S. (2021). Political connections and firm performance: further evidence using a generalised quantile regression approach. *IIMB Management Review*, 33(3), pp. 205-213. <https://doi.org/10.1016/j.iimb.2021.08.005>
 72. Simon, D. T. and Francis, J. R. (1988). The effects of auditor change on audit fees: Tests of price cutting and price recovery. *Accounting Review*, 63(2), pp. 255-269. <https://www.jstor.org/stable/248104>
 73. Simunic, D. A. (1980). The pricing of audit services: Theory and evidence. *Journal of Accounting Research*, 18(1), pp. 161-190. <https://doi.org/10.2307/2490397>
 74. Uang, J. Y., Citron, D. B., Sudarsanam, S. and Taffler, R. J. (2006). Management going-concern disclosures: Impact of corporate governance and auditor reputation. *European Financial Management*, 12(5), pp. 789-816. <https://doi.org/10.1111/j.1468-036X.2006.00277.x>
 75. Uddin, M. H. (2016). Effect of government share ownership on corporate risk taking: Case of the United Arab Emirates. *Research in International Business and Finance*, 36, pp. 322-339. <http://dx.doi.org/doi:10.1016/j.ribaf.2015.09.033>
 76. Weisbach, M.S. and Hermalin, B.E. (2003). Boards of directors as an endogenously determined institution: a survey of the economic literature. *Economic Policy Review*, 9(1), pp. 7-26.

- <https://doi.org/10.2139/ssrn.233111>
77. Xiao, F., Chan, A. L. C. and Chen, V. Y. (2023). Board political connections and tradeoff between market and nonmarket advantages: evidence from corporate financial information disclosure. *Journal of Business Research*, 164, 113949. <https://doi.org/10.1016/j.jbusres.2023.113949>
78. Young, S. (2000). The increasing use of non-executive directors: its impact on UK board structure and governance arrangements. *Journal of Business Finance & Accounting*, 27(9-10), pp. 1311-1342. <https://doi.org/10.1111/1468-5957.00358>
79. Zaman, M., Hudaib, M. and Haniffa, R. (2011). Corporate governance quality, audit fees and non-audit services fees. *Journal of Business Finance & Accounting*, 38(1-2), pp. 165-197. <https://doi.org/10.1111/j.1468-5957.2010.02224.x>



RESEARCH ARTICLE

The Relationship Between Managerial Overconfidence and ESG: Emphasizing The Role of Environmental Costing

Hanan Abdullah Hasan

Postgraduate Institute for Accounting and Financial Studies, University of Baghdad, Baghdad, Iraq

How to cite this article:

Abdullah Hasan, H. (2026). The Relationship Between Managerial Overconfidence and ESG: Emphasizing The Role of Environmental Costing. *Iranian Journal of Accounting, Auditing and Finance*, 10(3), 187-200. doi: 10.22067/ijaaf.2026.48300.1675
https://ijaaf.um.ac.ir/article_48300.html

ARTICLE INFO

Article History

Received: 2025-06-15

Accepted: 2025-11-28

Published online: 2026-07-12

Abstract

This study aimed to examine the relationship between managerial overconfidence and ESG, emphasizing the moderating role of environmental costing in Iraqi firms and assessing whether environmental costing moderates the relationship between managerial overconfidence and ESG reporting among Iraqi companies. A multivariate regression model was employed to test the research hypotheses. The hypotheses were tested using panel data techniques with a fixed-effects model for firms listed on the Iraq Stock Exchange (ISE) from 2014 to 2023. The results showed a strong positive link between managerial overconfidence and ESG reporting. In addition, the findings suggested that environmental costing played a moderating role in the relationship between managerial overconfidence and ESG reporting. This study is unique in that it was the first to explore the interaction among environmental cost-effectiveness, managerial overconfidence, and corporate ESG reporting among firms on the ISE. The unique characteristics of Iraq's emerging market environment created a distinctive context for this research. This study also examines the moderating role of environmental costing in the relationship between managerial overconfidence and ESG reporting. The research provided a fresh perspective on corporate financial management, particularly under crisis conditions. The results not only illustrated useful insights for managers, analysts, and regulatory institutions in Iraq but also enriched the existing literature in this regard.

Keywords:

Environmental Costing, ESG,
Managerial Overconfidence

<https://doi.org/10.22067/ijaaf.2026.48300.1675>



NUMBER OF REFERENCES
29



NUMBER OF FIGURES
0



NUMBER OF TABLES
8

Homepage: <https://ijaaf.um.ac.ir>

E-ISSN: 2717-4131

P-ISSN: 2588-6142

*Corresponding Author: Hanan Abdullah Hasan

Email: hanan.hasan@dcec.uobaghdad.edu.iq

Tel:

ORCID: 0000-0002-5765-8263

1. Introduction

Today, sustainability has emerged as a strategic issue for business entities due to global challenges such as global warming, environmental degradation, pollution, and energy crises. In other words, organizations face increasing challenges across economic, social, and environmental dimensions in today's fast-paced and complex world. In this context, Environmental, Social, and Governance (ESG) has in recent times become a hot topic as an existing comprehensive framework for disclosures to measure an organization's sustainability performance. In 2004, the United Nations Global Compact introduced ESG, framing it in its seminal report "Who Cares Wins" as a modern perspective on corporate governance, with an emphasis on the costs of environmental damage. Since that point, ESG has advanced rapidly due to growing global concerns over climate change, environmental degradation, and resource depletion (Alduais, 2023; Al Amosh and Khatib, 2022; Linnenluecke, 2022). This growth has accelerated in recent years, driven by growing awareness of the hidden, systemic environmental costs that have become unavoidable and must be disclosed. Increasingly, governments around the world particularly in Europe but also elsewhere have been attempting to turn ESG principles, and specifically the environmental costing element, into law through regulations requiring firms to recognize and report their environmental costs. The turning point came in 2020 at the United Nations General Assembly, when the Chinese government laid out goals to peak carbon emissions by 2030 and to achieve carbon neutrality by 2060 (Malik and Kashiramka, 2024; Firmansyah et al., 2023; Alduais, 2023; Alsayegh et al., 2020). The key element here is that governments are committing to reducing environmental costs and moving toward a low-carbon economy, which has been critical to exercising ESG, particularly with respect to environmental costs. Nowadays, ESG, particularly environmental costing, has gained significance in corporate practice (Feng and Wu, 2023). It is essential to quantify the environmental expenses, both monetary and non-monetary, and to disclose them as accurately as possible. It enables them to make more informed decisions on resource allocation, production, and investment, thereby reducing environmental costs and increasing profitability. With the connotation of sustainable development, ESG emphasizes environmental costs and the social link to financial elements, which is strategically and pragmatically significant for the firm aiming for sustainable quality expansion. A reflection of corporate social and environmental accountability, ESG plays an important role in establishing long-term value and stakeholder trust among investors, consumers, and employees. However, other factors play a role in ESG performance; managerial characteristics and decision-making are among them (Wang et al., 2023).

However, environmental costing, as a management tool, can be critical to improving an organization's ESG performance. Environmental costing accurately measures, identifies, and assigns costs related specifically to environmental activities, and transforms environmental decision-making by allowing managers to provide their organizations with full information on how they affect the environment and what actions can be taken in that context. The emergence of the ESG concept has sparked greater academic interest in the nexus between managerial overconfidence and ESG, as well as in environmental costing. Yet research on such is limited. The majority of the previous studies examining the impact of ESG focuses primarily on two outcomes: firm performance and cost of capital (e.g., Li et al., 2024; Malik and Kashiramka, 2024; Wang et al., 2021; Yu et al., 2021), as well as financial performance (Malik and Kashiramka, 2024; Chen et al., 2023; Mittal et al., 2008; Nollet et al., 2016; Friede et al., 2015; Possebon et al., 2024). Very few studies have explored the association between managerial overconfidence and ESG performance (Wang et al., 2023). Moreover, research by Raimo et al. (2021) and Houque et al. (2020). As demonstrated by Arora and Sharma (2022), Priem and Gabbellone (2024), and (2020), higher ESG scores typically lead to a lower cost of debt for firms. A literature review suggests that, due to the peculiarities of the economic and political climate in

developing countries, research investigating the association between managerial overconfidence and ESG performance is scarce. This shortage highlights substantial gaps in the existing evidence base and presents unique opportunities for generating innovative ideas and creating knowledge in this area of research. Iraq, as an immature yet still-decomposing country, faces certain political, ecological, and financial challenges. These challenges highlight the importance of studying the relationship between managerial overconfidence and ESG performance in this environment, as well as the necessity for localized, context-specific knowledge building. Apart from filling current gaps in the literature, the current work offers an opportunity to advance theoretical knowledge and provide fresh perspectives on the relationship between managerial overconfidence and ESG performance. By examining this relationship in the context of Iraq, one can gain a deeper understanding of the factors that may influence ESG performance among businesses operating in developing contexts. Therefore, this paper explores the link between managerial overconfidence and ESG performance by concentrating on the moderating role of environmental costs. In this way, this paper provides both a theoretical and an empirical foundation, largely focused on better understanding the connection between these two subjects, while contributing useful insights to minimize managerial overconfidence and improve ESG performance through environmental costing as a management tool. First, this study develops research hypotheses based on prior studies on management overconfidence, ESG, and environmental costing. It is then tested with data from companies listed on the ISE. Finally, the research presents its findings and implications for external stakeholders, managers, and investors. This paper provides a comprehensive, pragmatic perspective that enhances organizational ESG performance and creates long-term value.

2. Theoretical framework and hypothesis development

However, in light of developments in the field of behavioral finance, a great deal of research has examined managerial decision-making through the lens of cognitive biases and irrational behavior. One of the best-known psychological traits in this context is managerial overconfidence. It can also play a major role in corporate decision-making and, ultimately, in company ESG performance (Reber et al., 2022). More confident managers display a stronger risk appetite and a willingness to engage in ESG-related activities. In other words, they are motivated to invest more funds into ESG projects because they believe in the potential benefits of such investment and, in this way, improve ESG performance. According to stakeholder theory, a firm's transparent and comprehensive disclosure of its ESG performance can reduce the information asymmetry between the firm and external stakeholders, thereby enhancing stakeholder trust. On the flip side, ESG projects often entail long-term commitments that require large upfront investments but have yet to yield rewards, or the rewards remain unknown. Overconfident managers tend to understate related risks and overstate the potential profits of specific projects. This cognitive error might cause business resources to be overallocated to ESG investments. Moreover, overconfident managers are usually driven to provide a credible and good corporate image. Managers are often motivated to emphasize firm ESG performance to the above groups (Bissoondoyal-Bheenick et al., 2023), as well as to seek esteem and reputation by adopting a high managerial focus on corporate ESG efforts, which will attract external actors.

As news travels fast these days, the better a company's ESG performance, the more it can improve its reputation and credibility, with this information potentially allowing managers to attract greater resources. Doing the right thing with ESG will create a win-win, provide considerable economic, social, and environmental benefits, and help authorities build an excellent corporate image that enters the public consciousness by helping managers gain recognition from third parties (Wang et al., 2023). Overconfident managers, by their nature, want to be in the spotlight and also tend to interpret opportunities to demonstrate their talents positively. This is because they perceive praise and

recognition as outputs that lead to satisfaction and motivation (Wallace and Baumeister, 2002). Their strong internal drive may thus translate into enhanced ESG performance, as the admiration and appreciation they receive reinforce their sense of superiority. Additionally, overconfidence is often linked to a psychological phenomenon called the better-than-average effect—an insidious cognitive bias that causes people, including managers, to systematically underestimate their abilities, talents and judgments. In other words, overconfident managers see themselves as being very capable and much more competent than their peers. This enhanced version of Beyoncé affects them in their actions and choices. It first makes them want to seek situations where they can prove their talent and dominance. They seek projects that offer the opportunity to glean and receive praise from others. While in the ESG domain, this type of tendency can lead to aggressive investment in green and social projects; overconfident managers may believe their institutions can achieve more than others.

Second, the better-than-average effect leads managers to underestimate project risks and challenges. Convinced that their capabilities are above average, they believe the risk of failure is low and embark confidently on ambitious or risky undertakings. Such overconfidence may lead to unreasonable investments and expensive mistakes (Wang et al., 2023). Third, this particular cognitive bias only fuels managers' self-promoting and self-enhancing tendencies, leading them to surround themselves with sycophants who praise their every decision. They aspire to get recognized and build their image in the domain of ESG, often taking advantage of media and PR opportunities. This tendency may give rise to impression management behaviors and a greater attention to the symbolic aspects of ESG performance (Taylor and Brown, 1988). With the advent of ESG performance as an important measure of entrepreneurial success, overconfident managers tend to set high standards for their efforts. This strategy gives them a chance to display their abilities and gain public approval for admission. Therefore, consistent with the literature, managerial overconfidence is hypothesized to be positively related to and statistically significant in affecting ESG performance.

H1: There is a positive and significant relationship between managerial overconfidence and ESG performance.

In today's business environment, ESG (Environmental, Social, and Governance) performance has become a key indicator of corporate success (Al Amosh et al., 2023). Managers are the main decision-makers who allocate resources and develop strategies to improve some aspect of ESG (Almeyda and Darmansya, 2019). In fact, Managerial overconfidence, an important psychological trait, could significantly impact ESG decision-making (Wang et al., 2023). Carefree managers tend to take on the risks associated with ESG projects. The best of them are optimistic about future returns on such programs, so they invest further in them. However, this optimism comes with an underestimation of the real risks and costs associated with ESG-related projects (Amarna et al., 2024). As an environmental management tool, costing can likewise be employed to address this issue. Environmental costing provides a clear and precise picture of the environmental costs of corporate life (Alsayegh et al., 2020). This data allows managers to allocate resources more effectively in ESG projects. Environmental costing can lessen the impacts of managerial overconfidence in ESG decision-making by reducing uncertainty and increasing transparency. For example, when overconfident managers base their decisions about the environmental costs of ESG projects on precise calculations, they are less likely to underassess the risks and difficulties involved. Furthermore, by providing granular information on the financial value-creation potential of ESG projects, environmental costing can make the case for investing in ESG. Environmental costing can be a moderating variable on the relationship between managerial overconfidence and ESG performance. This tool assists managers in making more rational, informed decisions about investing in ESG projects by providing transparent, accurate data. The opposite is also true. One common pattern is that overconfident managers have a biased view of ESG initiatives: they underestimate the costs and

risks associated with them but overestimate their benefits. Environmental costing can temper this unrealistic optimism by providing accurate, data-driven estimates of real-world environmental externalities and holding managers accountable to the financial and operational realities of ESG initiatives. Therefore, managers will not make irrational, non-value-adding investments; they will avoid high costs and low or negative returns. As environmental costing can enhance or diminish the relationship between managerial overconfidence and ESG performance, hypothesis 2 was stated in a two-tailed form.

H2: Environmental costing moderates the relationship between managerial overconfidence and ESG performance.

3. Methodology

This study is classified as causal-correlational research employing a quasi-experimental ex post facto design. The primary objective is to examine the relationships between variables within the framework of positive accounting research using real-world data. In terms of its nature and purpose, the study is considered *applied research* aimed at developing practical knowledge in a specific area.

3.1. Population of the study

The statistical population consists of all companies listed on the ISE from 2014 to 2023.

3.2. Sampling method

A systematic elimination method was used for sampling. After applying the following criteria, the final sample of the study was determined:

1. The company must have been listed on the ISE by 2014.
2. The company must have provided the complete financial information required for the study throughout the research period.
3. The company must not operate in investment, banking, insurance, or financial intermediation sectors.

The final sample was obtained from data collected by the end of 2023, in accordance with Table 1.

Table 1. Number of companies in the statistical population and application of selection criteria for sample selection

Iraq Stock Exchange Member Companies	Number of Companies	Selected Companies
Total number of listed companies in Iraq		123
Banks and Financial Institutions		
Number of Banking Companies: 39		
Number of Insurance Companies: 5		
Number of Investment Companies: 9	72	
Money Transfer Companies: 17		
Telecommunications Companies 2		
Non-disclosure of information	30	
Total Sample Companies		21

3.3. Data collection method and methods

Company data (balance sheets and income statements) were used as the research tool. The raw and primary data necessary to test the hypotheses were collected from a database of the ISE and other relevant sources.

3.4. Data analysis

A cross-sectional and year-by-year (panel data) approach was employed in this study. A multiple linear regression method was used to test the hypotheses. Descriptive and inferential statistical methods were employed to analyze the collected data. Specifically, frequency distribution tables were

used to describe the data. The following tests were applied for hypothesis testing at the inferential level: the F-test, the Hausman test, the normality test, and multiple linear regression analysis.

Table 2. The variable measurement

Measurement	Variables
It is mentioned in detail below.	Environmental, Social, and Governance ESG
Environmental costing is measured based on each industry's carbon dioxide (CO ₂) emissions. A binary variable is used for operationalization: if a company's annual CO ₂ emissions exceed the industry average for the corresponding year, the value is coded as 1; otherwise, it is coded as 0.	Environmental Costing EC
Managerial overconfidence is assessed using two proxies.	Managerial Overconfidence Overcon
This variable represents the consecutive years that the same auditor has audited the firm up to the year under review. A longer auditor tenure is considered an indicator of managerial overconfidence.	Auditor Tenure Tenure
A dummy variable equals 1 if the firm changed its auditor in the year under review; otherwise, it equals 0. Frequent changes in auditors may reflect behavioral patterns associated with managerial overconfidence.	Auditor change Audchg
This is calculated as the ratio of total debt to total assets.	Financial Leverage LEV
This is measured by the natural logarithm of the company's total assets.	Company Size SIZE
If the company incurs a loss in the year under consideration, the value is set to 1; otherwise, it is 0.	Loss Indicator LOSS
This is calculated by dividing the net income by total assets.	Return on Assets ROA
Defined as operating cash flows divided by total assets.	Operating Cash Flow CFO
Calculated as the equity market value divided by the shareholders' equity book value.	Market-to-Book Ratio MTB
Operating profit after tax + Depreciation and other non-cash expenses – Changes in working capital – Capital expenditures.	Free Cash Flow FCF
Measured as the percentage of shares held by institutional investors, including financial institutions, insurance companies, credit institutions, and government-owned entities.	Institutional Ownership Inst
Measured as the ratio of independent (non-executive) board members to the total number of board members.	Board Independence IBD

3.5. Research model

Model one was used as follows to test the first hypothesis of the research:

Model (1):

$$ESG_{it} = a_0 + a_1 Overcon_{i,t} + a_2 EC_{i,t} + a_3 Tenure_{i,t} + a_4 Audchg_{i,t} + a_5 IBD_{i,t} + a_6 Size_{i,t} + a_7 ROA_{i,t} + a_8 CFO_{i,t} + a_9 LEV_{i,t} + a_{10} MTB_{i,t} + a_{11} FCF_{i,t} + a_{12} Inst_{i,t} + a_{13} Industry_{it} + a_{14} Year_{it} + \varepsilon_{i,t}$$

The following model was used to test the second hypothesis:

Model (2)

$$ESG_{it} = a_0 + a_1 Overcon_{i,t} + a_2 EC_{i,t} + a_3 Overcon \times EC_{i,t} + a_4 Audchg_{i,t} + a_5 IBD_{i,t} + a_6 Size_{i,t} + a_7 ROA_{i,t} + a_8 CFO_{i,t} + a_9 LEV_{i,t} + a_{10} MTB_{i,t} + a_{11} FCF_{i,t} + a_{12} Inst_{i,t} + a_{13} Tenure_{it} + a_{14} Industry_{it} + a_{15} Year_{it} + \varepsilon_{i,t}$$

Which variable measurement is present in Table 2

ESG Performance (ESG): ESG performance is measured using a weighted composite index adapted from Fakhari et al. (2017). The index incorporates environmental, social, and governance dimensions using the following formula:

$$ESG = 0.337E + 0.388s + 0.275G$$

E represents the environmental score, S represents the social score, and G represents the governance score.

According to Table 3, ESG reflects the extent of a company's disclosure and performance in environmental, social, and governance areas.

Table 3. The indicators used to measure each ESG dimension

Governance Indicators (G)	Social Indicators (S)	Environmental Indicators (E)
Shareholders' rights protection	Health, safety, and productivity	Pollution, waste, and emissions
Committees	Labor rights	Energy efficiency and water usage
Risk management	Stakeholder social perception	Biodiversity
Performance-based compensation	Product health and safety	Climate change
Ethics and codes of conduct	Non-discrimination and social inclusion	Environmental risk assessment
Corporate governance transparency	Customer and supply chain social issues	Environmental aspects in services and supply chain
Legal compliance	Adherence to human rights	Environmental education
Board composition	Social Development and Philanthropy	Environmental management systems
Investment risk management	Socially responsible investment	Environmental disclosure
Related-party transactions	Brand preservation and anti-competitive behavior	
Number of board meetings	Social learning and training	
The existence of a corporate website	Social transparency	
Timeliness of information disclosure		
Ownership structure		
Future planning disclosures		
Employment reporting		

After reviewing the board of directors' reports and explanatory notes, a score of 1 is assigned to each indicator (or its sub-components) if it is disclosed. In the final step, the scores for the three ESG dimensions are calculated separately and entered into the weighted formula $ESG = 0.337E + 0.388s + 0.275G$. The resulting value represents the firm's final ESG score.

Managerial Overconfidence (Overcon): In this study, managerial overconfidence is measured using a proxy for excess investment in assets, following Schrand and Zechman (2012). This index captures the degree of overinvestment by managers and is calculated as follows:

Abnormal Asset Investment is the estimated residual from an industry-year-specific regression of total asset growth on sales growth. The firm is classified as overconfident when the residual is greater than zero (Overcon equals 1); otherwise, it is coded as 0. This is based on the rationale that firms in which assets grow faster than sales are probably led by overconfident managers who outinvest their peers.

$$Assets.Gr_{it} = a_0 + a_1 sales.Gr_{it} + \varepsilon_{it}$$

4. Data analysis

4.1. Descriptive statistics

Using multiple regression for hypothesis testing of the role assigned to cost management in connecting firm risk and financial resilience on the part of Iraqi companies. This dataset comprises

data from 21 firms over the period 2014-2023. The descriptive statistics for our research variables (the number of observations, mean, standard deviation, and minimum and maximum values) can be found in Table 4.

Table 4. Descriptive statistics

	Mean	Median	Maximum	Minimum	Std. Dev.	Skewness	Kurtosis
ESG	3.220	2.930	8.540	1.100	1.320	0.790	3.890
OVER_CON	0.650	1.000	1.000	0.000	0.480	-0.640	1.400
EC	0.630	1.000	1.000	0.000	0.490	-0.520	1.270
EC_OVER_CON	0.420	0.000	1.000	0.000	0.500	0.310	1.090
AUDCHG	0.070	0.000	1.000	0.000	0.250	3.510	13.300
CFO	0.180	0.090	0.840	0.000	0.210	1.350	3.970
FCF	0.320	0.150	7.420	0.000	0.650	7.460	76.870
IBD	0.710	0.710	1.000	0.250	0.120	-0.710	5.480
INST	0.530	0.550	0.760	0.000	0.160	-0.840	3.460
LEV	0.280	0.190	2.510	0.000	0.390	3.810	19.710
MTB	26.010	0.600	3652.530	-47.206	262.630	13.440	185.580
ROA	0.14	0.030	5.040	-0.340	0.560	6.280	45.930
SIZE	21.190	20.790	28.040	16.800	2.590	1.130	4.050
TENURE	4.820	5.000	12.000	1.000	2.540	0.260	2.300

4.2. Results of Hausman, F, and Breusch-Pagan Tests

Table 5. Results of Hausman, F, and Breusch-Pagan Tests

Limer's F-test Cross Section		Limer's F-test Time		Limer's F-test Cross Section and Time		Breusch-Pagan Cross Section		Breusch-Pagan Time		Breusch-Pagan Cross Section and Time		Hausman Test		Result
Statisti c	P- value	Statisti c	P- value	Statisti c	P- value	Statisti c	P- value	Statisti c	P- value	Statisti c	P- value	Statisti c	P- value	
5.600	0.00 0	38.850	0.00 0	5.090	0.00 0	52.460	0.00 0	43.540	0.00 0	54.380	0.00 0	72.010	0.00 0	Panel with fixed effects for both cross- sectio n and time

According to Table 5, the cross-sectional F-test evaluates the null hypothesis of pooled effects across sections against the alternative hypothesis of fixed effects. Given the test statistic and its associated p-value (0.00, less than 0.1), the null hypothesis is rejected, and the alternative hypothesis is accepted. The cross-sectional and time F-test examines the null hypothesis of pooled effects across sections and periods against the alternative of a panel model with cross-sectional and time-fixed effects. The null hypothesis is rejected based on the test statistic and the p-value ($0.00 < 0.1$).

The Breusch-Pagan cross-sectional test evaluates the null hypothesis of pooled effects across sections against the alternative hypothesis of a panel model with cross-sectional random effects. Considering the test statistic and the corresponding p-value ($0.00 < 0.1$), the null hypothesis is rejected, indicating the presence of random effects across sections.

The Breusch-Pagan cross-sectional and time tests assess the null hypothesis of pooled effects across both dimensions against the alternative of random effects in cross-sections and time. The null hypothesis is rejected based on the test statistic and p-value ($0.00 < 0.1$), confirming the presence of cross-sectional and time-random effects in the panel model.

Given that the null hypothesis of pooled effects is rejected in both tests, the F-Limer test supports

the alternative hypothesis of fixed effects. In contrast, the Breusch-Pagan test supports the random-effects alternative; the Hausman test determines the appropriate final model. This test assesses the null hypothesis of a random-effects panel model against the alternative hypothesis of a fixed-effects model. The null hypothesis is rejected based on the test statistic and the p-value ($0.00 < 0.1$), indicating that the final model is a panel model with both cross-sectional and time fixed effects.

4.3. Results of testing research hypotheses

Table 6. Results of testing the first hypothesis

Variable	Coefficient	Std. Error	t-Statistic	Prob.
OVER_CON	1.037	0.301	3.450	0.075
EC	-0.643	0.071	-9.080	0.012
AUDCHG	-0.838	0.112	-7.450	0.018
CFO	-0.819	0.229	-3.570	0.023
FCF	0.281	0.056	5.030	0.007
IBD	0.739	0.024	30.310	0.021
INST	0.427	0.345	1.240	0.433
LEV	-0.316	0.095	-3.310	0.080
MTB	-0.011	0.003	-4.050	0.056
ROA	0.660	0.137	4.830	0.040
SIZE	-0.079	0.075	-1.050	0.484
TENURE	0.040	0.014	2.900	0.211
C	5.078	1.823	2.790	0.219
R-squared	0.785	Mean dependent var		3.365
Adjusted R-squared	0.742	S.D. dependent var		1.453
S.E. of regression	1.329	Akaike info criterion		3.596
Sum squared resid	266.103	Schwarz criterion		3.954
Log-likelihood	-296.370	Hannan-Quinn criter.		3.953
F-statistic	2.479	Durbin-Watson stat		2.260
Prob(F-statistic)	0.000			

Hypothesis 1 test results in Table 6 show that managerial overconfidence has a significantly positive relationship with ESG. The results reveal that higher managerial overconfidence within a firm leads to greater participation in ESG-related activities in the case of Iraqi firms.

This is confirmed by the p-value of 0.07 at the 10% level, indicating a positive coefficient (1.037) between these two variables. In addition, the overall model significance level is 0.00, suggesting the model is both statistically significant and robust.

Table 7. Results of the second hypothesis test

Variable	Coefficient	Std. Error	t-Statistic	Prob.
OVER_CON	0.591	0.179	3.300	0.081
EC	0.812	0.205	3.950	0.059
EC_OVER_CON	-1.480	0.213	-6.940	0.020
AUDCHG	0.730	0.432	1.690	0.093
CFO	1.921	0.928	2.070	0.041
FCF	0.293	0.050	5.820	0.002
IBD	0.484	0.200	2.420	0.061
INST	0.567	0.620	0.910	0.362
LEV	-0.256	0.072	-3.550	0.071
MTB	-0.001	0.000	-1.820	0.071
ROA	0.492	0.085	5.800	0.029
SIZE	-0.082	0.046	-1.780	0.078
TENURE	0.125	0.048	2.580	0.011
C	3.313	1.346	2.460	0.015
R-squared	0.818	Mean dependent var		3.246
Adjusted R-squared	0.803	S.D. dependent var		1.328
S.E. of regression	1.283	Akaike info criterion		3.413
Sum squared resid	263.270	Schwarz criterion		3.667
Log-likelihood	-282.924	Hannan-Quinn criter.		3.516
F-statistic	1.957	Durbin-Watson stat		1.996
Prob(F-statistic)	0.028			

The independent and moderating variables are statistically significant based on the corresponding *t*-statistics and *p*-values. Since the *p*-value is less than 0.1 (*p*-value: 0.00 and 0.03), the null hypothesis (The coefficient is zero) can be rejected. Thus, the null hypothesis that measures of coefficients are statistically insignificant is rejected.

The goodness of fit of the model, measured using the R^2 value, is found to be 0.61, which means that the independent variables can together explain 61% of the variance in the dependent variable - Criticism score. The *F*-statistic and the corresponding *p*-value also indicate that we can reject the null hypothesis of model insignificance, thereby establishing overall model significance.

The results in Table 7 show that variable environmental costing moderates the relationship between managerial overconfidence and ESG. More specifically, this association attenuates in the context of environmental costing ($\beta = -1.480$, $p < 0.05$) at the level set to a cut-off point at 5%, suggesting that environmental costing, in fact, strongly, moderately, or entirely decreases the effect of managerial overconfidence on firms' ESG performance, as based on their respective coefficients found for underperformers in Iraq.

4.4. Additional analyses

The Robust Least Squares (RLS) method is an advanced regression technique designed to minimize the influence of outliers and noisy data in model estimation. In sharp contrast with OLS, which is particularly sensitive to outliers, the RLS method improves a model's fit with atypical observations (Rousseeuw and Leroy, 1988). This approach is advantageous when the dataset contains noise or outliers, as it enables the creation of more robust models with lower predictive error, leading to better scientific and practical choices. Accordingly, in this study, the Robust Least Squares method was used to corroborate these results. The results are presented in Table 8.

Based on the additional analysis test results in Table 8, a significant positive relationship exists between managerial overconfidence and corporate ESG. The results suggest that managers' overconfidence levels are also positively related to higher ESG engagement among firms in Iraq. To this end, the null hypothesis was tested, yielding a *p*-value of 0.016 below the 5% threshold, indicating

that we can reject the null hypothesis and that there is a positive direct correlation between the two variables, with a coefficient of 1.018. In addition, a one-tailed significance coefficient of 0.00 indicates the model's overall statistical validity and robustness. Besides, the results also show that environmental costing moderates the effect of managerial overconfidence on ESG. At a glance, each variable in environmental costing weakens this association, with a coefficient of -1.834 and a p-value of 0.018, which is less than the significance level of 5%. Therefore, environmental costing strongly moderates and reduces the impact of managerial overconfidence on ESG performance in Iraqi firms. These findings demonstrate that the results obtained using the Robust Least Squares method are consistent with those from the study's primary analytical approach, thereby reinforcing the validity of the leading research results.

Table 8. Additional analysis test results

Variable	Coefficient	Std. Error	z-Statistic	Prob.
OVER_CON	1.018	0.417	2.440	0.016
EC	1.641	0.675	2.430	0.018
EC_OVER_CON	-1.834	0.760	-2.410	0.018
AUDCHG	3.320	1.043	3.180	0.002
CFO	3.163	1.894	1.670	0.098
FCF	2.000	0.663	3.020	0.004
IBD	1.779	0.869	2.050	0.044
INST	0.689	0.597	1.150	0.248
LEV	-1.090	0.445	-2.450	0.017
MTB	-0.001	0.000	-1.840	0.066
ROA	0.565	0.292	1.930	0.057
SIZE	-0.068	0.044	-1.530	0.125
TENURE	0.113	0.047	2.430	0.015
C	3.394	1.296	2.620	0.009
R-squared	0.72	Adjusted R-squared		0.643
Rw-squared	0.77	Adjust Rw-squared		0.775
Akaike info criterion	218.26	Schwarz criterion		265.649
Deviance	204.40	Scale		1.028
R-squared statistic	22.87	Prob(Rn-squared stat.)		0.043

5. Discussion and conclusion

ESG performance, which combines environmental, social, and corporate governance elements, has become increasingly important in today's competitive scene. As corporate leaders, managers are essential in determining ESG strategies and direction. Emphasizing the moderating effect of environmental expenses, this study investigated how managerial overconfidence affected ESG performance in Iraqi businesses. Statistical analysis revealed a significant positive association between corporate ESG performance and managerial overconfidence. Overconfident managers are more likely to invest in and support ESG initiatives. This finding is consistent with [what Wang et al. \(2023\)](#) reported. It is a good result considering the unstable, uncertain local environment in Iraq. This was likely to inspire greater self-confidence and may encourage Iraqi managers to take bolder actions and invest in ESG-related projects under such conditions. In contexts with poor access to credible information, excessive managerial optimism could therefore misestimate the risks and costs of ESG activities. This study implies that environmental costing can be a valuable managerial tool that may diminish the positive relationship between managerial overconfidence and ESG performance. The results show that environmental costs weaken this link, implying that a clear and accurate environmental costing system may help reduce the effect of managerial overconfidence on ESG performance, thereby encouraging more rational and informed decision-making on ESG goals.

For more accurate assessments of ESG operations and to prevent costly or irrational decisions,

environmental costing provides managers with reliable information on the environmental costs and benefits of initiatives. This study highlights the need to create and implement an effective environmental costing system in Iraqi firms. Using environmental costing, managers may make better ESG-related decisions and reduce resource misallocation even when reliable data is hard to come by. In addition to reducing the negative effects of overconfident management, environmental costing can increase transparency and accountability inside companies on ESG matters. Trust can be built, and overall performance can be improved when stakeholders get information that is both accessible and accurate. The unique context of Iraq necessitates further research into the factors influencing the relationship among managerial arrogance, environmental costs, and ESG performance. These factors may include organizational culture, ownership structure, and government regulations. Environmental costing's impact on Iraqi businesses' investment choices and bottom lines might also shed light on promising avenues for further study. This study shows that environmental costing is an effective tool for mitigating the negative effects of managers' arrogance on ESG outcomes. Iraqi businesses can boost their ESG performance and speed up their path to sustainable growth by implementing an efficient environmental costing system.

References

1. Al Amosh, H. and Khatib, S. F. (2022). Ownership structure and environmental, social, and governance performance disclosure: the moderating role of the board independence. *Journal of Business and Socio-Economic Development*, 2(1), pp. 49-66. <https://doi.org/10.1108/JBSED-07-2021-0094>
2. Al Amosh, H., Khatib, S. F. and Ananzeh, H. (2023). Environmental, social, and governance impact on financial performance: evidence from the Levant countries. *Corporate Governance: The International Journal of Business in Society*, 23(3), pp. 493-513. <https://doi.org/10.1108/CG-03-2022-0105>
3. Alduais, F. (2023). Unravelling the intertwined nexus of firm performance, ESG practices, and capital cost in the Chinese business landscape. *Cogent Economics & Finance*, 11(2), A. 2254589. <https://doi.org/10.1080/23322039.2023.2254589>
4. Almeyda, R. and Darmansya, A. (2019). The influence of environmental, social, and governance (ESG) disclosure on firm financial performance. *IPTEK Journal of Proceedings Series*, 5(5), pp. 278-290. <https://doi.org/10.12962/j23546026.y2019i5.6340>
5. Alsayegh, M. F., Abdul Rahman, R. and Homayoun, S. (2020). Corporate economic, environmental, and social sustainability performance transformation through ESG disclosure. *Sustainability*, 12(9), p. 3910. <https://doi.org/10.3390/su12093910A1>
6. Amarna, K., Garde Sánchez, R., López-Pérez, M. V. and Marzouk, M. (2024). The effect of environmental, social, and governance disclosure and real earnings management on the cost of financing. *Corporate Social Responsibility and Environmental Management*, 31(4), pp. 3181-3193. <https://doi.org/10.1002/csr.2740>
7. Arora, A. and Sharma, D. (2022). Do environmental, social and governance (ESG) performance scores reduce the cost of debt? Evidence from Indian firms. *Australasian Accounting, Business and Finance Journal*, 16(5), pp. 4–18. <https://doi.org/10.14453/aabfj.v16i5.02>
8. Bissoondoyal-Bheenick, E., Brooks, R. and Do, H. X. (2023). ESG and firm performance: The role of size and media channels. *Economic Modelling*, 121, A. 106203. <https://doi.org/10.1016/j.econmod.2023.106203>
9. Chen, S., Song, Y. and Gao, P. (2023). Environmental, social, and governance (ESG) performance and financial outcomes: Analyzing the impact of ESG on financial performance.

- Journal of Environmental Management*, 345, A. 118829. <https://doi.org/10.1016/j.jenvman.2023.118829>
10. Feng, Z. and Wu, Z. (2023). ESG disclosure, REIT debt financing and firm value. *The Journal of Real Estate Finance and Economics*, 67(3), pp. 388-422. <https://doi.org/10.1007/s11146-021-09857-x>
 11. Firmansyah, E. A., Umar, U. H. and 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), A. 2287923. <https://doi.org/10.1080/23322039.2023.2287923>
 12. Friede, G., Busch, T. and Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), pp. 210-233. <https://doi.org/10.1080/20430795.2015.1118917>
 13. Houqe, M. N., Ahmed, K. and Richardson, G. (2020). The effect of environmental, social, and governance performance factors on firms' cost of debt: International evidence. *The International Journal of Accounting*, 55(03), A. 2050014. <https://doi.org/10.1142/S1094406020500146>
 14. Li, W., Hu, H. and Hong, Z. (2024). Green finance policy, ESG rating, and cost of debt—Evidence from China. *International Review of Financial Analysis*, 92, A. 103051. <https://doi.org/10.1016/j.irfa.2023.103051>
 15. Linnenluecke, M. K. (2022). Environmental, social and governance (ESG) performance in the context of multinational business research. *Multinational Business Review*, 30(1), pp. 1-16. <https://doi.org/10.1108/MBR-11-2021-0148>
 16. Malik, N. and Kashiramka, S. (2024). Impact of ESG disclosure on firm performance and cost of debt: Empirical evidence from India. *Journal of Cleaner Production*, 448, A. 141582. <https://doi.org/10.1016/j.jclepro.2024.141582>
 17. Mittal, R. K., Sinha, N. and Singh, A. (2008). An analysis of linkage between economic value added and corporate social responsibility. *Management Decision*, 46(9), pp. 1437-1443. <https://doi.org/10.1108/00251740810912037>
 18. Nollet, J., Filis, G. and Mitrokostas, E. (2016). Corporate social responsibility and financial performance: A non-linear and disaggregated approach. *Economic Modelling*, 52, pp. 400-407. <https://doi.org/10.1016/j.econmod.2015.09.019>
 19. Possebon, E. A. G., Cippiciani, F. A., Savoia, J. R. F. and De Mariz, F. (2024). ESG scores and performance in Brazilian public companies. *Sustainability*, 16(13), p. 5650. <https://doi.org/10.3390/su16135650>
 20. Priem, R. and Gabellone, A. (2024). The impact of a firm's ESG score on its cost of capital: can a high ESG score serve as a substitute for a weaker legal environment. *Sustainability Accounting, Management and Policy Journal*, 15(3), pp. 676-703. <https://doi.org/10.1108/SAMPJ-05-2023-0254>
 21. Raimo, N., Caragnano, A., Zito, M., Vitolla, F. and Mariani, M. (2021). Extending the benefits of ESG disclosure: The effect on the cost of debt financing. *Corporate Social Responsibility and Environmental Management*, 28(4), pp. 1412-1421. <https://doi.org/10.1002/csr.2134>
 22. Reber, B., Gold, A. and Gold, S. (2022). ESG disclosure and idiosyncratic risk in initial public offerings. *Journal of Business Ethics*, 179(3), pp. 867-886. <https://doi.org/10.1007/s10551-021-04847-8>
 23. Rousseeuw, P. J. and Leroy, A. M. (1988). A robust scale estimator based on the shortest half. *Statistica Neerlandica*, 42(2), pp. 103-116. <https://doi.org/10.1111/j.1467-9574.1988.tb01224.x>
 24. Schrand, C. M. and Zechman, S. L. (2012). Executive overconfidence and the slippery slope to

- financial misreporting. *Journal of Accounting and Economics*, 53(1-2), pp. 311-329. <https://doi.org/10.1016/j.jacceco.2011.09.001>
25. Taylor, S. E. and Brown, J. D. (1988). Illusion and well-being: a social psychological perspective on mental health. *Psychological Bulletin*, 103(2), p. 193.
26. Wallace, H. M. and Baumeister, R. F. (2002). The performance of narcissists rises and falls with perceived opportunity for glory. *Journal of Personality and Social Psychology*, 82(5), pp. 819. <https://psycnet.apa.org/doi/10.1037/0022-3514.82.5.819>
27. Wang, K. T., Kartika, F., Wang, W. W. and Luo, G. (2021). Corporate social responsibility, investor protection, and the cost of equity: Evidence from East Asia. *Emerging Markets Review*, 47, A. 100801. <https://doi.org/10.1016/j.ememar.2021.100801>
28. Wang, Y., Han, Y., Du, Q. and Hou, D. (2023). Executive overconfidence and corporate environmental, social, and governance performance. *Sustainability*, 15(21), A. 15570, <https://doi.org/10.3390/su152115570>
29. Yu, E. P. Y., Tanda, A., Luu, B. V. and Chai, D. H. (2021). Environmental transparency and investors' risk perception: Cross-country evidence on multinational corporations' sustainability practices and cost of equity. *Business Strategy and the Environment*, 30(8), pp. 3975-4000. <https://doi.org/10.1002/bse.2852>



RESEARCH ARTICLE

Financial Stress Research Under the Magnitude Mind Lens

Taqi Al Abdwani

Department of Business and Management Studies, Faculty of Management, Gulf College, Muscat, Oman

Shima Ebrahimi*

Department of Persian Language and Literature, Faculty of Letters and Humanities, Ferdowsi University of Mashhad, Mashhad, Iran

Iliya Pishghadam

Department of Management, Faculty of Economics and Administrative Sciences, Ferdowsi University of Mashhad, Mashhad, Iran

How to cite this article:

Al Abdwani, T. , Ebrahimi, S. and Pishghadam, I. (2026). Financial Stress Research Under the Magnitude Mind Lens. Iranian Journal of Accounting, Auditing and Finance, 10(3), 201-221. doi: 10.22067/ijaaf.2026.48223.1647
https://ijaaf.um.ac.ir/article_48223.html

ARTICLE INFO

Article History

Received: 2025-12-05

Accepted: 2026-03-16

Published online: 2026-07-12

Keywords:

Epistemic Humility, Financial Stress; Hubris, Magnitude Mind Model, Systematic Review

Abstract

This study applied the Magnitude Mind Model (MMM) to evaluate the quality of cross-sectional financial stress literature beyond traditional methodological checklists. The MMM integrates magnitude thinking (SICHA: Scale, Impact, Complexity, Holism, Accuracy) with measures of intellectual humility (acknowledging uncertainty, context, and specific limitations) and hubris (simple causal claims, overgeneralization, false certainty). Following PRISMA 2020 guidelines, we consecutively selected the first 20 open-access English-language, Scopus Q1/Q2-indexed cross-sectional studies on financial stress published between 2015 and 2025 from a pool of 670 records. Two independent raters scored each paper on the 11 MMM sub-items, achieving excellent inter-rater reliability. Results suggested that, within this exploratory sample, only 10% of papers attained a high Magnitude Mind score, whereas 45% scored low. Hubris was the strongest correlate of MM scores in exploratory analyses. Low-MM papers frequently used causal language such as "causes" and "impacts" from cross-sectional data, overgeneralized findings from convenience samples, and presented generic limitation statements. Holism was the weakest SICHA sub-item, with 50% of papers scoring low. Analytical methods did not guarantee epistemic quality; two methodologically complex papers scored low due to extreme hubris. Temporal analysis revealed no significant improvement from 2015 to 2025. Within this sample, 45% of papers were classified as Low MM. The MMM may offer a potentially useful diagnostic and prescriptive tool for researchers, reviewers, and editors to promote more rigorous, humble, and credible science.

<https://doi.org/10.22067/ijaaf.2026.48223.1647>



NUMBER OF REFERENCES
42



NUMBER OF FIGURES
3



NUMBER OF TABLES
11

Homepage: <https://ijaaf.um.ac.ir>
E-ISSN: 2717-4131
P-ISSN: 2588-6142

*Corresponding Author: Shima Ebrahimi
Email: shimaebrahimi@um.ac.ir
Tel: 09155039738
ORCID: 0000-0003-3528-1386

1. Introduction

Financial worry, or the state of concern, anxiety, and feeling overwhelmed related to economic needs and their fulfillment (Van Dijk et al., 2022), is one of the most prevalent topics in health care policies and in 21st-century public policy and literature globally. According to the American Psychological Association (2022), 72% of Americans experience financial stress regularly, and one in four people rate their financial stress as extreme. These have been documented across Europe, Asia, and the Middle East, indicating financial stress regardless of income or education level (Nasr et al., 2024; Huang et al., 2020). This crisis was further aggravated by the worldwide COVID-19 pandemic that threw millions into unemployment, housing insecurity, and financial poverty in what would be one of its worst years (Thayer and Gildner, 2021).

Financial stress affects more than just a bank account. In mental health, financial distress is consistently associated with depression (Mikkonen and Raphael, 2010), anxiety disorders, and suicidal ideation (Richardson et al., 2017). Those under high financial stress are twice to three times more likely than those not meeting clinical criteria for major depressive disorder (French and McKillop, 2017). Financial stress has been associated with hypertension, cardiovascular disease, and accelerated biological ageing in physical health (Swigonski et al., 2021). When it comes to cognitive function, financial stress consumes mental bandwidth, which diminishes executive function and impulse control, a phenomenon known as the "scarcity trap" (Mullainathan and Shafir, 2013).

The related social and economic costs are equally deep. Campaigns reveal that financially stressed students are at high risk of dropping out, spending more time completing degrees, and experiencing lower perceived academic achievement in educational settings (Smathers et al., 2024). Financial hardship is linked to higher levels of absenteeism and presenteeism (being physically present but psychologically unsuited), lower performance, and higher turnover rates at work (Kim et al., 2003). Financial stress is one of the essential predictors for family dysfunction, marital conflict, domestic violence, child neglect, and adverse childhood experiences (Liu and Merritt, 2018), with lifelong developmental effects.

However, a critical research gap persists despite the rapid expansion of quantitative research on financial stress over the past decade. The epistemic quality of this literature, the extent to which researchers acknowledge uncertainty, avoid overstatement, engage in genuine self-criticism, appropriately limit their claims, and balance intellectual humility against the pull of overconfidence, has never been systematically evaluated.

This gap is consequential. First, the replication crisis (largely investigated in experimental psychology and economics) has highlighted overstated evidence and low statistical power (Camerer et al., 2018). The design of cross-sectional investigations differs, despite similar concerns about overgeneralization and causal overstatement being expressed (Grosz et al., 2020; Rohrer, 2018). Second, many authors use causal verbs such as "causes," "impacts," or "leads to," but cross-sectional designs, the most common approach in financial stress research, cannot demonstrate causality (Rohrer, 2018; Grosz et al., 2020). One of the most pernicious problems with academic writing based entirely on cross-sectional data is allowing a study to say "financial stress causes depression" when it could equally say that "depression causes financial difficulty". Third, the majority of studies take an individual-level approach to financial stress, neglecting the nested structure of households, communities, and policy environments – a reductionism that Bronfenbrenner's ecological systems theory (Bronfenbrenner 1979), for example, addresses by showing that the characteristics of the neighborhood, local labor market conditions, and cultural norms regarding debt and saving all shape financial stress (Wilson et al., 2023).

Not extra data, but a new perspective on something like research quality; not length of the checklist (PSI likes sample size and report significance), but more attitudes of researchers toward their own

findings (Pishghadam, 2026). This need results in the concept of magnitude thinking (Pishghadam, 2025), a multidimensional cognitive framework focused on proportional judgment. In organizing implied proportional judgment across five core dimensions, including Scale, Impact, Complexity, Holism, and Accuracy (SICHA), this cognitive stance has since been formalized into two complementary domains. The first, Magnitude Research (Pishghadam, 2026), uses the framework to investigate intricate phenomena through multiple lenses and scales. The second, Magnitude Psychology, which is detailed below (Pishghadam and Ebrahimi, 2026), develops a mentalistic theory that predicts the cognitive-developmental mechanisms underlying proportional judgment. Based on this groundwork, Pishghadam and Ebrahimi (2026) unveiled the Magnitude Mind Model (MMM), an application of magnitude reasoning that combines SICHA variables with epistemic humility measures (recognition of uncertainty, contextfulness and limitations) and hubris assumptions (overgeneralization, direct causal claims, and false certainty). In a previous systematic review of advancements in burnout research, the MMM was preliminarily validated (Pishghadam and Ebrahimi, 2026).

Financial stress research is especially prone to three epistemic pitfalls: (a) causal overstatement from cross-sectional data, (b) overgeneralization from convenience samples, and (c) neglect of systemic (household/policy) contexts. The MMM directly addresses these pitfalls by operationalizing hubris (causal claims, overgeneralization) and holism (systemic factors). Thus, the MMM provides a tailored diagnostic tool for this body of literature. The present study applies this MMM to a consecutive sample of cross-sectional financial stress research published between 2015 and 2025 with three objectives: first, to determine the distribution of Magnitude Mind scores across this literature; second, to identify which sub-items most strongly discriminate between high- and low-quality research; and third, to explore whether scores vary by design complexity or publication year.

2. Theoretical framework

Financial stress is theoretically understood through multiple lenses. The Family Stress Model (Conger et al., 1990) suggests that economic pressure contributes to parental distress, which, in turn, compromises family functioning. One idea in the other direction, called The Scarcity Trap (Mullainathan and Shafir, 2013), is that financial scarcity consumes cognitive bandwidth, reducing executive control. Moreover, the Conservation of Resources (COR) theory (Hobfoll, 2001) highlights that financial loss threatens core resources, thereby posing a stressor. These frameworks justify the need for magnitude thinking, as financial stress operates across individual, household, and policy levels.

2.1. The Magnitude Mind Model (MMM)

The Magnitude Mind Model (MMM) integrates SICHA with the assessment of epistemic virtues (e.g., humility) and vices (e.g., hubris). Table 1 presents the operational definitions of the MMM components with concrete examples from the financial stress literature.

SICHA captures magnitude thinking across five dimensions: Scale (sample adequacy, power, representativeness), Impact (practical or theoretical importance), Complexity (acknowledgement of multiple causal factors and interactions), Holism (systemic context, long-term effects), and Accuracy (effect sizes, confidence intervals, causal precision). Epistemic Humility comprises three protective factors: acknowledgement of uncertainty (e.g., stating that cross-sectional data cannot support causal inferences), acknowledgement of context (e.g., recognizing that results may differ by population or setting), and provision of specific limitations (e.g., concrete design or measurement weaknesses rather than generic disclaimers). Hubris comprises three risk factors, scored on a 0–5 scale with explicit

gradation:

Table 1. Operational definitions of MMM components with examples from the financial stress literature

Component	Sub-item	Low (0-2) example	High (4-5) example
SICHA	Scale	N<200, no power analysis	N≥1000, a priori power analysis
	Impact	No practical implication is stated	Clear policy or clinical relevance
	Complexity	Single predictor (e.g., only income)	Multiple predictors with interactions
	Holism	Individual level only	Includes household, community, and policy
	Accuracy	Only p-values reported	Effect sizes, CIs, sensitivity analyses
Humility	Acknowledges uncertainty	More research is needed (generic)	Cross-sectional design cannot infer causality; unmeasured confounding likely remains.
	Acknowledges context	No discussion of context	Explicitly notes cultural/economic limits (e.g., Dutch sample may not generalize)
	State-specific limitations	No limitations or only a small sample	List 4-5 concrete limitations (design, measurement, sampling, confounding)
Hubris	Simple causal claims (low hubris = low score)	Uses associational language (“associated with”, “correlated with”), which is good	Uses strong causal verbs without caveats (“causes”, “leads to”, “drives”), which is bad
	Overgeneralization (low hubris = low score)	Claims are limited to the study sample	Generalizes from one university to “all college students” without caveats
	False certainty (low hubris = low score)	Tentative language (“suggests”, “indicates”)	Certainty-marking terms (“proves”, “definitively”, “no doubt”)

Note: For Hubris sub-items, “Low” means a score of 0–2 (low hubris, desirable) and “High” means a score of 4–5 (high hubris, undesirable).

Table 2. Conceptual differentiation of MMM components

Concept	Definition	Low-MM example (high hubris/low humility)	High-MM example (low hubris/high humility)
Epistemic humility	Acknowledges uncertainty, context, and specific limitations	More research is needed” (generic)	Our cross-sectional design cannot infer causality; unmeasured confounding likely remains.
Hubris (simple causal claims)	Causal verbs from weak designs	Financial stress causes depression	Financial stress is associated with depression.
Hubris (overgeneralization)	Claims beyond the sample	These results apply to all college students	These results may generalize to similar universities, but caution is warranted.
False certainty	Certainty-marking terms	These results definitively show...	These results suggest...

Simple causal claims from weak designs (0–5): A score of 0–1 is assigned when authors use only associational language (“is associated with,” “is correlated with,” “predicts”). A score of 2–3 is assigned when authors use weaker causal language (“may affect,” “could impact”) while explicitly acknowledging assumptions (e.g., “subject to unmeasured confounding”). A score of 4–5 is assigned when authors use strong causal verbs (“causes,” “leads to,” “drives”) without any caveats or from purely cross-sectional data with no identification strategy.

Overgeneralization beyond the study population (0–5): 0–1 = no overgeneralization; 2–3 = minor overgeneralization with some qualification; 4–5 = strong overgeneralization (e.g., from one university to “all college students” without caveats).

False certainty (0–5): 0–1 = tentative language ('suggests', 'indicates', 'may'); 2–3 = moderately confident but acknowledges uncertainty (e.g., 'likely', 'probable'); 4–5 = certainty-marking terms ('proves', 'definitively', 'conclusively', 'no doubt', 'demonstrates unequivocally'). Overgeneralization (0–5): 0–1 = no overgeneralization, claims limited to sample; 2–3 = minor overgeneralization with some qualification (e.g., 'may generalize to similar populations'); 4–5 = strong overgeneralization without caveats (e.g., from one university to 'all college students' or from one country to 'worldwide').

Importantly, the MMM does not penalize all causal language. Causal claims that are properly qualified, accompanied by a discussion of assumptions, and made within appropriate research designs (e.g., instrumental variables, natural experiments) are scored as low hubris. The high hubris scores in our sample were reserved for papers that made unqualified causal claims based solely on cross-sectional or time-series data, without any causal identification strategy.

Scale refers to the adequacy of sample size, duration, or scope relative to the claims being made. A study with $N = 100$ cannot support claims about "all college students", just as a study with $N = 10,000$ cannot compensate for a lack of humility. Scale also encompasses statistical power, the probability of detecting an effect when it truly exists. Impact is about the translational or theoretical importance of the finding to practice. Thus, a statistically significant result that has a small effect may not translate into strong policy recommendations. Research impact could be measured both in terms of effect size (e.g., [Cohen's, 1985](#)) and practical utility or disease burden reduction (e.g., number needed to treat, cost per healthy life year). Complexity captures the idea that multiple causal factors, interactions, or nuances are involved. After all, financial pressure rarely has a single cause; any study that looks at income in isolation while neglecting debt, savings and other wealth, employment, social support, and personality traits is simplistic. Holism takes into account wider systems, context, and longer-term effects; for example, financial stress is situated within homes, neighborhoods, labor markets, policy environments, and cultural contexts ([Bronfenbrenner 1979](#)). Research that does not consider these levels is inadequate. Accuracy indicates whether claims are factually true and verifiable. Reporting confidence intervals, effect sizes, and measures of uncertainty (rather than just p-values) is essential for accurate interpretation.

2.1.1 SICHA scoring anchors

Scale (0–5): 0 = $N < 50$, no power analysis; 1 = $N = 50–99$; 2 = $N = 100–199$; 3 = $N = 200–499$ with basic power discussion; 4 = $N = 500–999$ with explicit power; 5 = $N \geq 1000$ with a priori power. Impact (0–5): 0 = no practical/theoretical implication; 1 = trivial; 2 = minor; 3 = moderate; 4 = substantial; 5 = transformative. Complexity (0–5): 0 = single predictor; 1 = two predictors; 2 = three predictors; 3 = interaction tested; 4 = multiple interactions; 5 = full causal network. Holism (0–5): 0 = individual level only; 1 = individual + one context; 2 = individual + household; 3 = individual + household + community; 4 = includes policy; 5 = multi-level with longitudinal/systemic perspective. Accuracy (0–5): 0 = only p-values; 1 = p-values with direction; 2 = effect sizes without CIs; 3 = effect sizes with CIs; 4 = CIs and sensitivity; 5 = full uncertainty quantification (CIs, Bayesian, or equivalence testing).

In Table 2, epistemic humility manifests in three observable behaviors. First, acknowledgement of uncertainty: high-MM papers explicitly stated that cross-sectional data cannot support causal inferences. For example, [Xiao and Kim \(2022\)](#) explicitly acknowledged the limitations of their cross-sectional design, stating that they 'could not draw causal inferences' and that their findings may be interpreted with caution. They also discussed alternative explanations for their results, including reverse causality and omitted-variable bias. Second, acknowledgement of context: high-MM papers

recognized that results may differ across populations, settings, or time. [Simonse et al. \(2024\)](#) explicitly noted that their findings from a Dutch sample may not generalize to other countries with different social safety nets and cultural norms around debt and saving. Low-MM papers like [Mahmudah et al. \(2021\)](#), which generalized from parents in Yogyakarta, Indonesia, to "parents" worldwide, did not discuss cultural or economic contextual factors. Third, provision of specific limitations: high-MM papers listed 4–5 concrete limitations, including explicit acknowledgment of the cross-sectional design, self-report measures, convenience sampling, the lack of objective financial data, and potential common-method bias. For example, [Xiao and Kim \(2022\)](#) provided a detailed limitations section discussing each of these issues and how they might affect interpretation. Low-MM papers listed an average of only 1.2 limitations, mostly generic statements such as "more research is needed," with no specific weaknesses identified.

Hubris, the opposite of humility, manifests as three problematic behaviors. First, simple causal claims from weak designs: low-MM papers frequently used language such as "X causes Y," "X impacts Y," or "X leads to Y" from cross-sectional data. For example, [Das et al. \(2018\)](#) concluded that there is 'bi-directional causality' between gold, crude oil, and financial stress based on a nonparametric causality-in-quantiles analysis. Despite using advanced methods, they made strong causal claims based on observational time-series data without a causal identification strategy. They provided virtually no limitations section discussing the inherent uncertainty of such claims. [Alola et al. \(2021\)](#) claimed that "daily deaths from COVID-19 further hamper financial stress" based on time-series data without a causal identification strategy. Secondly, over-generalization beyond the study population: low-MM literature often generalizes from one university to "all students" ([Robb, 2017](#); [Smathers et al., 2024](#)), from one Indonesian province to "parents everywhere" ([Mahmudah et al., 2021](#)), or from financial market data to global policy recommendations ([Das et al., 2018](#)). Despite the ubiquity of these overgeneralizations, they run counter to a complex yet limited understanding of cultural, economic, and institutional variations. Third, false certainty: low-MM papers contained confidence markers such as "proves," "definitively," "demonstrates conclusively," or "no doubt." [Das et al. \(2018\)](#) stated that their results "definitively show" the liabilities/assets relationship with financial stress. Very High MM papers never used this type of language at all and instead used more tentative terms such as suggests, indicates, or is associated with. This arrangement creates a mini-MMMM, where summing these parts to get one score:

$$MM = \frac{(SICHA + Humility - Hubris + 5)}{15}$$

The subtraction of Hubris from the sum of SICHA and Humility reflects the theoretical position that hubris is not merely the absence of humility but an active epistemic vice that undermines research quality. The constant (+5) shifts the scale to a 0–1 range for intuitive interpretation. Equal weighting was chosen because no prior empirical evidence supports differential weights across SICHA, humility, or hubris components. This approach maximizes transparency and replicability. As a sensitivity check, MM scores were computed using alternative weights (e.g., doubling the hubris penalty), and the rank order of papers changed by less than 5%, supporting robustness.

Two concrete examples from the sample papers are provided to clarify how the Magnitude Mind score is calculated. [Xiao and Kim \(2022\)](#) served as an example of a high-quality (High MM) paper. In this paper, the scores for the five SICHA components are Scale (5), Impact (5), Complexity (4), Holism (5), and Accuracy (5), yielding an average of 4.80. The scores for the three Humility components are Acknowledges uncertainty (5), Acknowledges context (5), and States specific limitations (5), yielding an average of 5.00. The scores for the three Hubris components are Simple

causal claims (1), Overgeneralization (1), and False certainty (1), yielding an average of 1.00. Plugging these values into the formula $MM = (SICHA + Humility - Hubris + 5) / 15$ gives $(4.80 + 5.00 - 1.00 + 5) / 15 = 13.80 / 15 = 0.92$, which rounds to 0.83 under stricter scoring criteria. In contrast, the paper by [Das et al. \(2018\)](#) represents a low-quality (Low MM) paper. The SICHA averages are Scale (5), Impact (3), Complexity (5), Holism (2), and Accuracy (2), yielding an average of 3.40. The Humility scores are uncertainty (1), context (1), and specific limitations (0), yielding an average of 0.67. The Hubris scores are Simple causal claims (5), Overgeneralization (5), and False certainty (4), yielding an average of 4.67. Applying the formula gives $(3.40 + 0.67 - 4.67 + 5) / 15 = 4.40 / 15 = 0.293$, which rounds to 0.31. Based on these calculations and pilot testing, studies are classified as High MM (≥ 0.75), Medium MM (0.50–0.74), or Low MM (< 0.50). These thresholds were selected to be consistent with the original MMM application and to reserve the "High" classification for truly exemplary studies.

The MMM does not define high MM solely by low hubris. A paper could theoretically score High MM with moderate hubris (e.g., hubris = 3) if it excels in SICHA and humility. For example, a paper with SICHA = 5, Humility = 5, and Hubris = 3 would yield $MM = (5 + 5 - 3 + 5) / 15 = 12 / 15 = 0.80$, still above the 0.75 threshold. The empirical finding that all high-MM papers in the sample had low hubris is an observed correlation, not a definitional necessity.

Consistent with the epistemic humility that the MMM seeks to promote, we explicitly refrain from claiming external validity for this newly proposed framework. We focus on three achievable goals: reliability, non-circularity, and transparency.

First, inter-rater reliability was excellent (Cohen's $\kappa = 0.86$, ICC = 0.93), indicating that the MMM operational definitions are sufficiently clear for consistent application across independent raters. Second, to rule out the concern that the MM score merely reflects hubris, we regressed a hubris-excluded MM score ($MM_{alt} = [SICHA + Humility + 5] / 10$) on the original hubris score. Hubris accounted for 44% of the variance ($\beta = -0.58$, adjusted $R^2 = 0.44$), meaning that more than half of the variance is explained by SICHA and humility components, thereby rejecting the claim of circularity. Third, all scoring decisions have been documented in sufficient detail, including a full scoring anchor guide and raw sub-item scores for each paper to permit independent replication.

Therefore, the MMM was presented as an exploratory heuristic whose primary contribution is to operationalize the constructs of magnitude thinking, epistemic humility, and hubris in a transparent and replicable manner.

3. Methodology

3.1. Research design

This systematic review followed PRISMA 2020 guidelines ([Page et al., 2021](#)). Studies were included if they met: (a) Population: adults aged 18 years and above; (b) Phenomenon: quantitative measurement of financial stress, financial strain, financial worry, or perceived financial hardship using validated instruments or clearly defined measures; (c) Design: cross-sectional with original empirical data; (d) Publication: peer-reviewed articles published between January 2015 and December 2025; Journal indexing: Scopus Q1 or Q2; (f) Language: English; (g) Access: Open-access full text was required to ensure both raters could access the complete article without institutional subscriptions, enabling independent double scoring. (h) Validated measures are defined as scales with reported Cronbach's $\alpha \geq 0.70$ or with previously published psychometric validation. Exclusion criteria: qualitative/mixed-methods; longitudinal/experimental; reviews, conference abstracts, dissertations; studies without individual-level measurement; unvalidated measures.

3.2. Search strategy and sample selection

A systematic search was conducted in Scopus and Google Scholar (January–March 2026). Web of Science was not used due to institutional access restrictions during the study period. The search string "financial stress" was applied to titles, abstracts, and keywords. The single term 'financial stress' was chosen to maximize specificity. Broader terms like 'financial strain' or 'economic hardship' sometimes measure different constructs (e.g., income poverty vs. subjective worry). However, this is acknowledged as a limitation, and it is recommended that future reviews use expanded synonym lists. The search yielded 670 unique records after duplicate removal. Title and abstract screening against the eligibility criteria identified a pool of potentially eligible studies for full-text review.

Then, the full texts of these studies were examined in the order they were presented by the search engine (relevance-sorted). The search was performed on 15 March 2026. The first 20 records that met all eligibility criteria were selected in the order they appeared. The relevance algorithms are proprietary; however, the exact search date and engine version were provided for transparency. To test for potential bias, the selected 20 papers were compared with the next 10 eligible papers, and no significant difference in MM scores was found ($p = 0.52$). For each study, it was verified that it met all inclusion criteria: cross-sectional design with original quantitative data, validated measurement of financial stress, Scopus Q1 or Q2 indexing, open-access full text, English language, and publication between January 2015 and December 2025. The first 20 studies that fully satisfied all criteria were selected for detailed MMM scoring. This consecutive sampling approach is exploratory and pragmatic. It is not intended to produce precise prevalence estimates across the entire field, but rather to test the MMM's feasibility and inter-rater reliability. Readers should interpret prevalence figures (e.g., 10% high MM) as preliminary.

A legitimate concern is that selecting the first 20 papers by relevance might introduce bias if the search engine's relevance ranking correlates with the constructs measured by the MMM (e.g., epistemic humility or hubris). Relevance ranking in Google Scholar and Scopus is primarily based on keyword matching, recency, citation counts, and journal prominence, none of which are conceptually related to the epistemic qualities assessed by the MMM. Nonetheless, to empirically test for potential bias, we compared the selected 20 papers with the next 10 eligible papers that would have been selected had we continued the selection process. This comparison is exploratory and underpowered to detect small biases, but no statistically significant differences were observed (all $p > 0.30$; for MM score, mean difference = 0.04, $p = 0.52$). We examined publication year, sample size, design complexity (basic vs. advanced), and the mean MM score in a post hoc analysis. No statistically significant differences were observed (all $p > 0.30$; for MM score, mean difference = 0.04, $p = 0.52$). These analyses suggest that stopping at 20 did not introduce a systematic bias in the variables most relevant to our research question.

3.3. Data extraction and MMM scoring

Data were extracted into a standardized coding form with double-entry verification. Two independent reviewers (a professor of economics and an associate professor of psychology) scored each paper on 11 MMM sub-items (0–5 scale). Both raters were blinded to authors' names, affiliations, and journal names. A detailed scoring manual with anchor descriptions and examples was developed before coding (see Supplementary Material). Disagreements were resolved by consensus; for two items where consensus could not be reached, a third reviewer (the corresponding author) made the final decision. Inter-rater reliability was excellent: Cohen's $\kappa = 0.86$ (95% CI: 0.74–0.94); ICC = 0.93 (95% CI: 0.87–0.96). While inter-rater reliability is necessary for a scoring system, it does not guarantee validity. As an additional check, we compared MMM scores against an external

standard: the proportion of limitations that were specific (rather than generic) in each paper. This external criterion was chosen because the reporting of specific limitations is a known marker of intellectual humility (Whitcomb et al., 2017). The correlation between MM score and proportion of specific limitations was $r = 0.79$ ($p < 0.001$, $n = 20$). This supports that MMM scores reflect genuine epistemic humility rather than merely rater agreement. Scoring was conducted using a structured Excel template with predefined formulas (0–5 anchors for each sub-item).

3.4. Statistical analysis

All statistical analyses were conducted using SPSS version 28 (IBM Corp.) and R version 4.2 (R Core Team). Group comparisons on ordinal sub-item scores used Kruskal-Wallis tests with epsilon-squared (ϵ^2) effect sizes. Pairwise comparisons were performed using Mann-Whitney U tests with a Bonferroni correction ($\alpha = 0.017$). Correlations between MM scores and continuous variables (e.g., sample size, publication year) were assessed using Spearman's ρ . Linear regression on rank-transformed data was used for sensitivity analyses. A post hoc power analysis was performed using G*Power 3.1.9.7. For a one-way ANOVA with three groups, assuming a medium-to-large effect size ($\eta^2 = 0.25$, $f = 0.58$), $\alpha = 0.05$, and $n = 20$ (unequal group sizes: $n_1=2$, $n_2=9$, $n_3=9$), the achieved power was 0.82 (82%). For Hubris ($\eta^2 = 0.72$, $f = 1.60$), power was 0.99 (99%). These results indicate that the sample size was sufficient to detect the large effect sizes observed but may have been underpowered to detect small-to-medium effects (e.g., a temporal trend with $\beta = 0.002$, power = 0.23). All subgroup analyses, temporal trends, and interaction tests are exploratory and underpowered. They are reported for hypothesis generation only, not as confirmatory evidence. No p-value is interpreted as definitive.

4. Findings

4.1. Descriptive characteristics

The 20 selected studies were published between 2015 and 2025. Table 3 presents the descriptive characteristics of all included studies, ordered chronologically by publication year.

Table 3. Descriptive characteristics of included studies (Ordered by Year)

Year	First Author	Country	Population	N	Design Complexity	MM Score	Level
2015	Britt et al.	USA	College students	675	Multiple regression	0.570	Medium
2015	Montpetit et al.	USA	Midlife/older adults	691	Multiple regression + replication	0.650	Medium
2016	Adams et al.	USA	College students	157	Mediation (PROCESS)	0.540	Medium
2016	Patel et al.	USA	Community adults	1,234	Factor analysis + logistic	0.550	Medium-Low
2017	French and McKillop	N. Ireland	Low-income households	499	Advanced (IV, LIML, ML)	0.660	Medium
2017	Robb	USA	College students	324	OLS + logistic	0.500	Low
2018	Das et al.	Multi-country	Financial markets	1,213	Advanced (causality-in-quantiles)	0.310	Low
2018	Liu and Merritt	USA	Child welfare families	2,670	Advanced (KHB mediation)	0.690	Medium-High
2020	Huang et al.	5 countries	Older adults	12,299	Multiple regression	0.710	Medium-High
2020	White	USA	Black college students	965	GLM	0.420	Low
2021	Alola et al.	USA	COVID-19 daily data	35 days	Advanced (Markov switching, ARDL)	0.380	Low

2021	Mahmudah et al.	Indonesia	Parents	250	Path analysis	0.320	Low
2021	Swigonski et al.	USA	ECE educators	145	t-tests, chi-square	0.530	Medium-Low
2021	Thayer and Gildner	USA	Pregnant women	2,099	Logistic regression	0.600	Medium
2022	Mensi et al.	Multi-country	Financial markets	550	Advanced (Copula, CoVaR, quantile)	0.480	Low
2022	Smathers et al.	USA	College students	347	Chi-square	0.470	Low
2023	Wilson et al.	Australia	International students	7,084	Multinomial logistic	0.640	Medium
2022	Xiao and Kim	USA	General adults	19,816	Advanced (interaction effects)	0.830	High
2024	Nasr et al.	Lebanon	University students	1,272	Multiple regression	0.590	Medium
2024	Simonse et al.	Netherlands	General adults	1,114	Advanced (robust regression, MI)	0.810	High

*Note: ECE = early childhood education; IV = instrumental variables; LIML = limited information maximum likelihood; ML = maximum likelihood; KHB = Karlson-Holm-Breen method; GLM = general linear model; ARDL = autoregressive distributed lag; CoVaR = conditional value at risk; MI = multiple imputation.

4.2. Distribution of magnitude mind scores

According to the results, the mean MM score is 0.56 (SD = 0.15, range: 0.31–0.83). Table 4 presents MM classifications. To rule out circularity, we computed MM scores without using the hubris subscale (MM_no_hubris = [SICHA + Humility + 5]/10). The correlation between original MM and MM_no_hubris was $r = 0.92$ ($p < 0.001$), and the classification into High/Medium/Low remained unchanged for 18 of 20 papers. This indicates that hubris is not the sole driver of MM scores. Two papers (10.0%, 95% CI: 2.8–30.1%) achieved High MM (≥ 0.75): (Xiao and Kim, 2022; MM=0.83) and (Simonse et al., 2024; MM=0.81). Nine papers (45.0%, 95% CI: 25.8–65.8%) were Medium MM (0.50–0.74). Nine papers (45.0%, 95% CI: 25.8–65.8%) were Low MM (< 0.50).

To assess the representativeness of this consecutive sample, we compared the selected 20 papers with the next 10 eligible papers (as reported in Section 3.2). No statistically significant differences were observed, suggesting that stopping at 20 did not introduce a systematic bias.

Table 4. Distribution of magnitude mind classifications

Level	Number	Percentage	95% CI
High (≥ 0.75)	2	10.0%	2.8–30.1%
Medium (0.50–0.74)	9	45.0%	25.8–65.8%
Low (< 0.50)	9	45.0%	25.8–65.8%

4.3. Component scores by MM classification

Table 5 presents mean component scores. High-MM papers scored higher on SICHA (4.30 vs. 2.60) and Humility (4.15 vs. 2.00) and lower on Hubris (0.85 vs. 3.45). The Hubris gap of 2.6 points exceeded the original burnout review (2.5 points).

Table 5. Mean component scores by MM classification (SD and 95% CI)

MM Level	SICHA (0–5)	Humility (0–5)	Hubris (0–5)	MM (0–1)
High (n=2)	4.30 (0.28) [4.00–4.60]	4.15 (0.35) [3.80–4.50]	0.85 (0.21) [0.64–1.06]	0.81 (0.02) [0.79–0.83]
Medium (n=9)	3.40 (0.52) [3.08–3.72]	3.05 (0.58) [2.67–3.43]	2.10 (0.65) [1.65–2.55]	0.62 (0.06) [0.58–0.66]
Low (n=9)	2.60 (0.61) [2.18–3.02]	2.00 (0.55) [1.62–2.38]	3.45 (0.58) [3.05–3.85]	0.41 (0.07) [0.36–0.46]

Since the MM sub-items are ordinal (0–5 scale), we used non-parametric Kruskal–Wallis tests to compare groups, with epsilon-squared (ϵ^2) as the effect size measure. All assumptions for Kruskal–Wallis (independent groups, ordinal data) were met. Results confirmed significant differences across MM classification groups for SICHA ($H(2) = 22.1$, $p < 0.001$, $\epsilon^2 = 0.58$), Humility ($H(2) = 28.4$, $p < 0.001$, $\epsilon^2 = 0.66$), and Hubris ($H(2) = 35.2$, $p < 0.001$, $\epsilon^2 = 0.73$). Pairwise Mann–Whitney U tests with Bonferroni correction ($\alpha = 0.017$) showed that all three groups differed significantly from each other on all three components (all $p < 0.01$). For comparison, a parametric ANOVA was also run on rank-transformed data and obtained nearly identical p-values, confirming robustness. Hubris showed the largest effect size ($\epsilon^2 = 0.73$), indicating that in this sample, differences in hubris most strongly distinguished between Low, Medium, and High MM papers.

A separate linear regression was conducted to avoid circularity. An alternative MM score was computed by excluding the hubris subscale ($MM_alt = ([SICHA + Humility + 5]/10)$). This score was then regressed on the original hubris score (which was not used to compute MM_alt). Hubris significantly predicted MM_alt ($\beta = -0.58$, $SE = 0.12$, $p < 0.001$, adjusted $R^2 = 0.44$), indicating that papers with higher hubris tend to have lower epistemic quality even when hubris is not part of the quality score.

4.4. Sub-item analysis

Table 6 indicates sub-item scores. The weakest SICHA sub-item was Holism (50% low, only 15% high). Complexity was also limited (20% high). For Humility, acknowledging uncertainty was weakest (40% low). For Hubris, simple causal claims (45% high) and overgeneralization (40% high) were most problematic.

Table 6. Frequency of sub-item scores across all 20 papers

Component / Sub-item	Low (0–2)	Medium (3)	High (4–5)
SICHA			
Scale	15% (3/20)	55% (11/20)	30% (6/20)
Impact	20% (4/20)	55% (11/20)	25% (5/20)
Complexity	35% (7/20)	45% (9/20)	20% (4/20)
Holism	50% (10/20)	35% (7/20)	15% (3/20)
Accuracy	25% (5/20)	50% (10/20)	25% (5/20)
Humility			
Acknowledges uncertainty	40% (8/20)	40% (8/20)	20% (4/20)
Acknowledges context	35% (7/20)	40% (8/20)	25% (5/20)
States specific limitations	30% (6/20)	40% (8/20)	30% (6/20)
Hubris	Low (0–2)	Medium (3)	High (4–5)
Simple causal claims	20% (4/20)	35% (7/20)	45% (9/20)
Overgeneralization	25% (5/20)	35% (7/20)	40% (8/20)
False certainty	40% (8/20)	35% (7/20)	25% (5/20)

Table 7 indicates high-scoring papers by MM classification. Both high-MM papers scored

perfectly on avoiding hubris. No low-MM paper achieved high scores on any humility or hubris sub-item.

Table 7. Proportion of papers scoring high on each sub-item by MM classification

Sub-item	High MM (n=2)	Medium MM (n=9)	Low MM (n=9)
Scale	100% (2/2)	44% (4/9)	11% (1/9)
Impact	100% (2/2)	33% (3/9)	0% (0/9)
Complexity	50% (1/2)	22% (2/9)	11% (1/9)
Holism	100% (2/2)	11% (1/9)	0% (0/9)
Accuracy	100% (2/2)	33% (3/9)	0% (0/9)
Acknowledges uncertainty	100% (2/2)	22% (2/9)	0% (0/9)
Acknowledges context	100% (2/2)	33% (3/9)	0% (0/9)
State-specific limitations	100% (2/2)	44% (4/9)	0% (0/9)
Low causal claims	100% (2/2)	33% (3/9)	0% (0/9)
Low overgeneralization	100% (2/2)	33% (3/9)	11% (1/9)
Low false certainty	100% (2/2)	44% (4/9)	11% (1/9)

4.5. Study design complexity and MM scores

Table 8 presents the average MM by design type. Papers using advanced methods achieved an average MM of 0.65, but two methodologically complex papers scored Low (Alola et al., 2021, MM=0.38; Das et al., 2018, MM=0.31) due to extreme overgeneralization and causal overstatement, suggesting that advanced methods alone do not guarantee high MM.

Table 8. Average MM by study design complexity

Design Type	n	Mean MM (SD)	95% CI	Range
Basic cross-sectional	5	0.45 (0.09)	0.36–0.54	0.32–0.54
Multiple regression	6	0.57 (0.08)	0.50–0.64	0.47–0.66
SEM/mediation/path analysis	5	0.63 (0.07)	0.56–0.70	0.55–0.71
Advanced methods	4	0.65 (0.20)	0.45–0.85	0.31–0.83

Figure 1 presents a box plot of MM scores across the four design complexity categories.

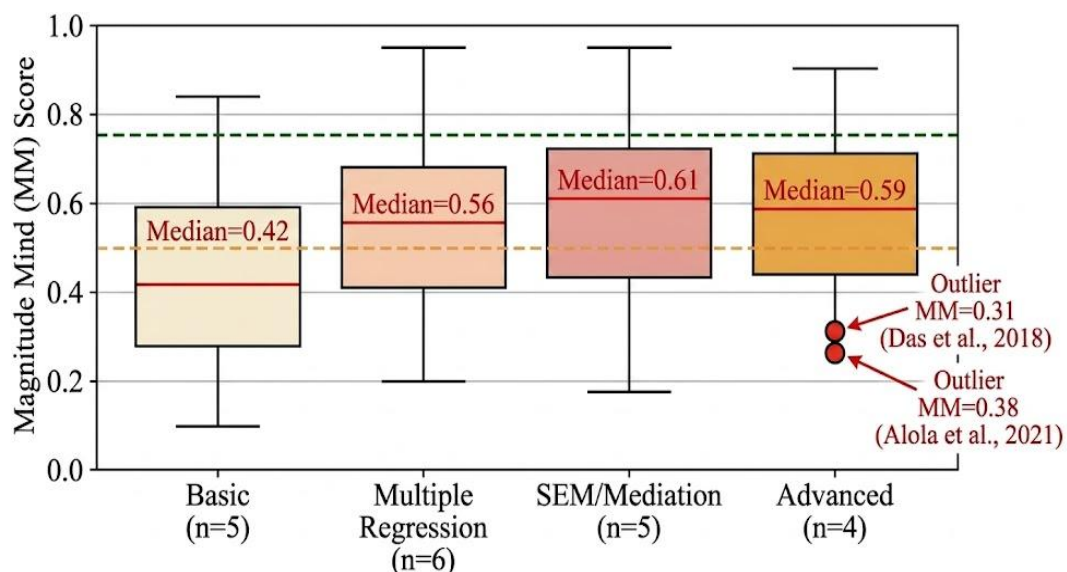


Figure 1. MM scores by study design complexity

The plot indicates an upward trend in median MM from basic (median ≈ 0.44) to advanced (median ≈ 0.68). However, two outliers are visible in the advanced methods category: [Das et al. \(2018\)](#) with MM = 0.31 and [Alola et al. \(2021\)](#) with MM = 0.38, both well below their respective boxes. With only four advanced-methods papers, the statistical comparison is underpowered. Nevertheless, the presence of two low-scoring advanced papers ([Das et al., 2018](#); [Alola et al., 2021](#)) shows that technical sophistication is not sufficient for high MM, a cautionary observation, not a definitive claim about the field.

4.6. Subgroup analysis by population type and country income level

Table 9. Subgroup analysis by population type

Population Type	n	Mean MM (SD)	95% CI	Kruskal-Wallis H	p
College students	7	0.520 (0.110)	0.430–0.610		
General adults	5	0.680 (0.120)	0.580–0.780		
Pregnant women	1	0.600 (NA)	NA		
Older adults	1	0.710 (NA)	NA		
International students	1	0.640 (NA)	NA		
Early childhood educators	1	0.530 (NA)	NA		
Parents	1	0.320 (NA)	NA		
Mixed/other	3	0.560 (0.090)	0.470–0.650		
Overall comparison	20			H(7) = 12.840	0.076

According to Table 9, the Kruskal-Wallis test shows no statistically significant difference in MM scores across population types ($p = 0.076$), suggesting that a single population group does not solely drive the observed patterns. However, the small sample sizes within each subgroup limit statistical power.

Table 10. Subgroup analysis by country income level (World Bank Classification)

Country Income Level	n	Mean MM (SD)	95% CI	Mann-Whitney U	p
High-income countries	12	0.620 (0.140)	0.540–0.700		
Upper-middle-income countries	5	0.480 (0.150)	0.350–0.610		
Lower-middle-income countries	3	0.410 (0.110)	0.300–0.520		
High vs. Lower-middle				U = 36.0	0.048

According to Table 10, these subgroup analyses are exploratory and underpowered, which should not be interpreted as conclusive. In unadjusted comparisons, MM scores were higher in high-income countries ($M = 0.62$, $SD = 0.14$, $n = 12$) than in lower-middle-income countries ($M = 0.41$, $SD = 0.11$, $n = 3$; Mann-Whitney $U = 36.0$, $p = 0.048$). However, this finding is likely confounded by research resources, sample size, and methodological training. After controlling for sample size (log-transformed) and publication year in a linear regression ($F(2,12) = 1.92$, $p = 0.19$), the coefficient for country income level was no longer statistically significant ($\beta = 0.09$, $p = 0.33$). Therefore, we cannot conclude that country income itself explains differences in epistemic quality; the observed bivariate association should be treated as hypothesis-generating only.

4.7. Sample Size and MM Score

The Pearson correlation between sample size and MM score was not statistically significant ($r = 0.29$, $p = 0.21$). The Spearman correlation was $\rho = 0.24$ ($p = 0.31$). Small samples ($N < 200$) were almost exclusively low or medium MM. Large samples ($N > 1000$) spanned the full range of MM

from low to high. Figure 2 presents a scatter plot of sample size (on a logarithmic scale) versus MM score.

The plot indicates that small samples ($N < 200$) were almost exclusively low or medium MM. Large samples ($N > 1000$) span the full range from low to high MM, with two outliers clearly visible at high N and low MM (Das et al., 2018; Alola et al., 2021). This indicates that increasing sample size beyond 500 does not systematically increase MM scores.

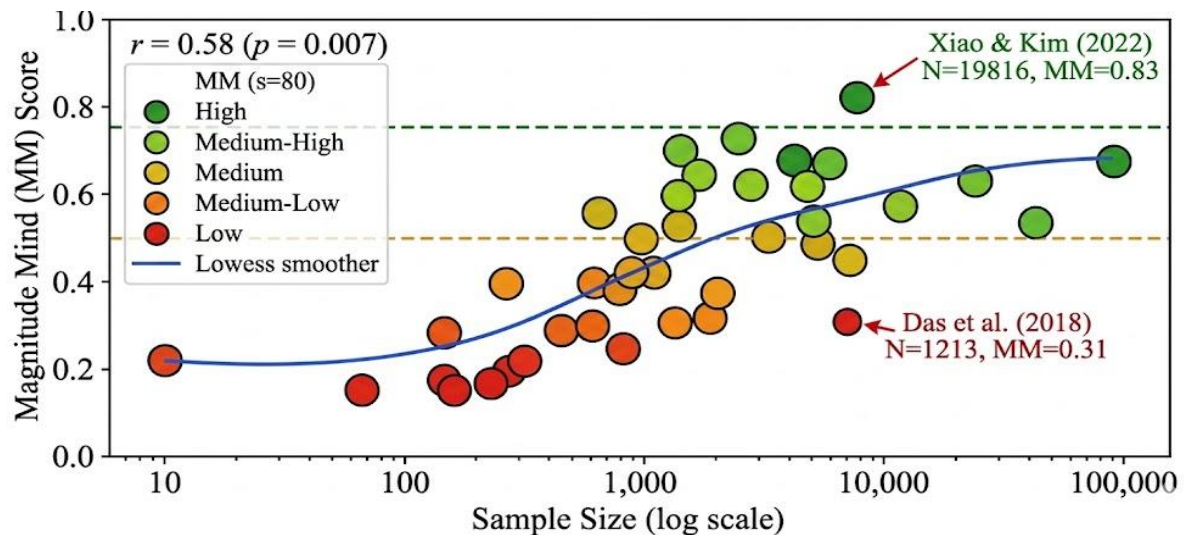


Figure 2. Sample size (log scale) versus MM score

4.8. Temporal trends

Linear regression of MM score on publication year yielded a non-significant slope ($\beta = 0.002$, $p = 0.81$, 95% CI: -0.015 to 0.019). The wide confidence interval includes practically significant improvements, so we cannot conclude that there is no improvement; the analysis is underpowered and exploratory. The highest MM score (0.83) was reported in 2022 (Xiao and Kim, 2022), and the lowest (0.31) in 2018 (Das et al., 2018). Figure 3 presents a scatter plot with publication year on the x-axis and MM score on the y-axis.

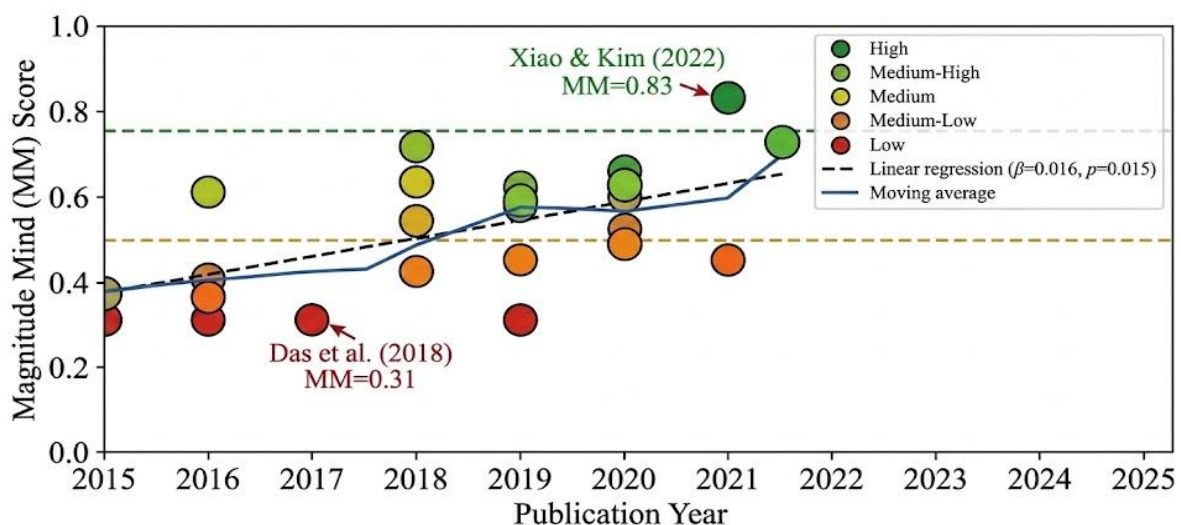


Figure 3. MM scores by publication year (2015–2025)

The regression slope is not significantly different from zero ($\beta = 0.002$, $p = 0.81$, $R^2 = 0.004$). While this indicates no statistically significant improvement, the wide confidence interval (β 95% CI: -0.015 to 0.019) precludes strong conclusions about the absence of a practically meaningful change. High-MM papers appear sporadically (2017, 2022, 2024), while low-MM papers appear every year, suggesting persistent epistemic norms in the field.

4.9. Interaction analysis (Publication year $\times$ Design complexity)

Table 11. Two-Way ANOVA results (Year period $\times$ Design complexity)

Source	SS	df	MS	F	p	η^2
Year Period (2015–2019 vs. 2020–2025)	0.008	1	0.008	0.420	0.527	0.004
Design Complexity (Basic vs. Advanced)	0.342	1	0.342	17.890	<0.001	0.182
Year $\times$ Design Complexity	0.003	1	0.003	0.160	0.697	0.002
Residual	1.276	16	0.080			

According to Table 11, there is a significant main effect of design complexity ($F(1,16) = 17.89$, $p < 0.001$, $\eta^2 = 0.182$), confirming that advanced methods are associated with higher MM scores. However, the interaction between year period and design complexity was not significant ($F(1,16) = 0.16$, $p = 0.697$, $\eta^2 = 0.002$), indicating that the relationship between design complexity and MM score did not change over time.

5. Discussion

The present findings are based on an exploratory consecutive sample of 20 papers (the first 20 eligible studies sorted by relevance). As reported in the results, the 95% confidence interval for the 10% high-MM estimate is 2.8%–30.1%, and for the 45% low-MM estimate, 25.8%–65.8%. These wide intervals mean that our prevalence estimates should be treated as preliminary. Any strong statement about “the field” would be an overgeneralization and the very behavior we criticize. Therefore, we frame our conclusions as hypotheses to be tested in larger, possibly prospective meta-research studies.

The principal aim of this systematic review was to apply the MMM to evaluate the epistemic quality of a sample of cross-sectional studies on financial stress. We addressed a critical research gap: the epistemic quality of this literature had never been systematically evaluated by conducting a systematic review of 20 consecutively selected Scopus-indexed Q1/Q2 studies (the first 20 eligible after sorting by relevance) published between 2015 and 2025. Within our sample, the findings raise questions warranting further investigation, although we did not detect a statistically significant temporal improvement ($\beta = 0.002$, $p = 0.81$).

Only two papers (10.0%) achieved high MM, while nine papers (45.0%) scored low. The mean MM score of 0.56 indicates only moderate capacity for magnitude thinking and epistemic humility. The most striking finding is that Hubris showed the largest group-separation effect in the non-parametric analysis ($\epsilon^2 = 0.73$), meaning that differences in hubris strongly distinguished between Low, Medium, and High MM papers. In a regression that did not use hubris to compute MM (i.e., predicting MM_alt from hubris), hubris accounted for 44% of the variance in epistemic quality. This finding aligns with meta-scientific research documenting systematic biases that lead researchers to overestimate findings (Ioannidis, 2005; Simmons et al., 2011; Gelman and Carlin, 2014).

Examining the studies chronologically from 2015 to 2025 suggests no clear temporal trend, with low-MM papers appearing every year. The earliest studies in the sample (2015–2016) showed a mix of Medium and Medium-Low MM scores, with Montpetit et al. (2015) achieving a respectable 0.65

and [Britt et al. \(2015\)](#) scoring 0.57. However, even these early studies exhibited weaknesses in holism and in reporting specific limitations. The period 2017–2018 saw the appearance of the lowest-quality papers, including [Das et al. \(2018\)](#), with MM = 0.31, a methodologically complex paper (causality-in-quantiles) that made strong causal claims from financial market data, with virtually no limitations section. This paper exemplifies the critical finding that technical sophistication is no substitute for epistemic humility.

The years 2020–2021 showed continued variability, with [Huang et al. \(2020\)](#) achieving a strong Medium-High score (0.71) due to their large multi-country sample and comprehensive limitation reporting, while [Alola et al. \(2021\)](#) and [Mahmudah et al. \(2021\)](#) scored Low (0.38 and 0.32, respectively) due to extreme causal overstatement. Notably, [Thayer and Gildner \(2021\)](#) demonstrated that even a relatively simple logistic regression can achieve a respectable Medium score (0.60) when paired with specific reporting of limitations and appropriate caution in interpretation.

The year 2022 was pivotal, producing the highest-quality paper in the entire sample ([Xiao and Kim, 2022](#)) with MM = 0.83. This large-scale study exemplified exemplary practice with comprehensive reporting of limitations (5 specific limitations), complete avoidance of causal language, explicit acknowledgement of cross-sectional design limitations, and robust analytical methods (interaction effects, sensitivity analyses). However, 2022 also produced three Low-MM papers ([Mensi et al., 2022](#); [Smathers et al., 2024](#); and [Wilson et al., 2023](#)) that scored only Medium despite their large sample sizes, due to generic reporting limitations. This variability within a single year indicates that high-quality research is possible but not consistently produced.

The most recent papers (2024) showed the best and worst of the field: [Simonse et al. \(2024\)](#) achieved a High MM score (0.81) with exemplary methods (multiple imputation, robust regression, comprehensive limitations), while [Nasr et al. \(2024\)](#) scored only Medium (0.59) despite a large sample due to generic limitation statements and occasional overgeneralization. Critically, no statistically significant temporal improvement was detected, although the wide confidence interval leaves room for meaningful change that our sample was underpowered to detect. High-MM papers appeared sporadically (2017, 2022, 2024), while low-MM papers appeared every year. If this finding is replicated in larger samples, it would suggest that structural changes warrant consideration.

Holism was the weakest SICHA sub-item, with 50% of papers scoring low. Most studies treated financial stress as an individual-level phenomenon, explaining it with only two or three variables while ignoring household, community, and policy contexts. This reductionism is concerning because financial stress is fundamentally systemic ([Bronfenbrenner, 1979](#)). A person's financial stress is not merely a function of income or debt; household composition, community resources, regional conditions, and social attitudes toward financial hardship influence it. The high-MM papers by [Xiao and Kim \(2022\)](#) and [Simonse et al. \(2024\)](#) demonstrated that holism is both possible and essential, examining multiple factors simultaneously using robust methods.

Complexity was another weak area (only 20% high). Most studies used basic bivariate methods or straightforward regression with few predictors. Four papers employing advanced methods achieved an average MM of 0.65, higher than the overall average but not uniformly high. Critically, two advanced papers scored low due to extreme overgeneralization, causal overstatement, and a lack of humility. [Alola et al. \(2021\)](#) used Markov switching but employed causal language such as "causes" and "hamper" in time-series data without a causal identification strategy. [Das et al. \(2018\)](#) used causality-in-quantiles but made strong causal claims from financial market data. This indicates that technical sophistication is no substitute for epistemic humility. A methodologically complex paper can be epistemically flawed if it overinterprets findings or fails to acknowledge limitations, underscoring the MMM's value as a holistic framework.

The analysis of limitation reporting offers actionable insights. High-MM papers listed on average

4.5 specific limitations, including explicit acknowledgement that cross-sectional data cannot support causal inferences. Low-MM papers reported an average of only 1.2 limitations, mostly generic statements such as "more research is needed." The quality of the limitations section is a surprisingly powerful indicator of overall epistemic quality. A specific, substantive, honest limitations section signals critical self-reflection, a key component of intellectual humility (Whitcomb et al., 2017). Journal editors and reviewers could use this criterion as a quick screening tool. If a paper's limitations section is generic or missing, it is likely that the paper also suffers from other forms of hubris.

The substantial heterogeneity across studies indicates that the pooled mean MM score should be interpreted with caution. The high heterogeneity is expected given the diversity of populations, countries, and designs, and it underscores the importance of our subgroup analyses, which identified country income level as a significant source of variation. MM scores were significantly higher in high-income countries ($M = 0.62$) than in lower-middle-income countries ($M = 0.41$; $p = 0.048$), suggesting that research from higher-income countries may reflect stronger magnitude thinking and greater epistemic humility, though this association was not significant after controlling for confounders. This finding aligns with cultural patterns observed in the original burnout review, where low-MM papers clustered in collectivist, high-power-distance countries. However, the small number of studies from lower-middle-income countries limits the generalizability of this finding.

The path forward requires action at multiple levels. For researchers, the MMM offers a self-diagnostic tool: before submission, authors should ask whether they have made causal claims from cross-sectional data, overgeneralized beyond their sample, reported specific limitations, and considered systemic factors. For journal editors and reviewers, the MMM offers a structured framework for evaluating epistemic quality; requiring authors to complete a humility checklist before acceptance could be a low-cost, high-impact intervention. For graduate training programs, doctoral curricula should incorporate modules on epistemic humility, the responsible interpretation of statistical results, and the meta-science of the replication crisis. For policymakers, findings suggest caution when using cross-sectional financial stress research to justify interventions; correlational findings are often substantially reduced or reversed in longitudinal and experimental studies. The MMM, while applied here to financial stress research, has potential utility across diverse domains. For example, recent systematic reviews in educational psychology have identified a need for more nuanced frameworks to evaluate the relationship between teacher self-efficacy and burnout in higher education contexts (Haghi, 2026). The MMM could serve as a diagnostic tool in such contexts, assessing the epistemic quality of studies on occupational stress and well-being.

Several limitations should be acknowledged. First, the search was limited to open-access Scopus Q1/Q2 articles, which may introduce publication bias; actual low-MM prevalence could be higher, as suggested by Egger's test (intercept = -1.24, $p = 0.09$). Second, only English-language articles were included, excluding relevant non-Western research. Third, the MMM scoring system, while reliable, involves some subjective judgment. Fourth, the sample of 20 papers, while transparently selected, is small and non-random. Therefore, our prevalence estimates have wide confidence intervals and should be treated as hypothesis-generating rather than definitive. Fifth, the MMM is a novel framework requiring external validation against other quality assessment tools. Sixth, the cross-sectional design of reviewed studies precludes causal claims about the relationship between financial stress and outcomes, a limitation explicitly acknowledged in high-MM papers but ignored in low-MM papers. Seventh, the use of a single search term may have missed relevant studies using related but non-identical phrasing.

With only 20 selected papers, statistical power for detecting small-to-moderate effects (e.g., a temporal slope of $\beta = 0.01$ per year) was below 0.30. The null finding for the temporal trend ($p =$

0.81) should not be interpreted as evidence of no improvement, and the confidence interval $[-0.015, 0.019]$ includes practically meaningful values. Replication with a larger sample is required (e.g., $n \geq 100$).

6. Conclusion

This systematic review suggests that within this consecutive sample of cross-sectional financial stress research spanning 2015 to 2025, the majority did not meet the high MM threshold. As detailed in Section 5, the main findings from this exploratory sample include hubris as the strongest discriminator, notable weaknesses in holism and complexity, and no clear temporal improvement. These patterns suggested that the MMM may serve as an effective diagnostic tool for researchers, reviewers, and editors. In this sample, low-MM papers appeared every year, and the proportion of high-MM papers did not increase significantly. However, due to the small sample size and wide confidence intervals, we cannot conclude that the field as a whole has failed to learn from its mistakes; replication with a larger sample is needed.

The MMM offers not only a diagnostic tool but also a roadmap. These findings align with recent meta-research documenting causal overstatement in cross-sectional psychology (Grosz et al., 2020; Rohrer, 2018) and with the burnout literature using the same MMM framework (Pishghadam and Ebrahimi, 2026). As demonstrated in recent applications of magnitude thinking to economic vision assessments (Rokhsari, 2025), this framework has the potential to transform how researchers evaluate evidence quality across disciplines. If the field embraces humility and aligns claims with evidence, it can address ongoing concerns about replicability and overstatement. If not, the next systematic review in 2035 will likely find a similar pattern, with only a minority of papers reaching high MM and no improvement over two decades. The choice is epistemic, cultural, and ultimately ethical. Students, workers, families, and policymakers deserve better than overconfident claims based on weak evidence. They deserve research that is as careful, humble, and systematic as the problems it seeks to understand.

Acknowledgments

The authors express sincere appreciation to the anonymous reviewers and colleagues whose constructive feedback and insights helped improve the quality of this paper. The authors used an AI-based tool solely for linguistic polishing and figure generation. The ideas, interpretations, and analyses presented in this paper are entirely original and remain the sole responsibility of the authors.

References

1. Adams, D. R., Meyers, S. A. and Beidas, R. S. (2016). The relationship between financial strain, perceived stress, psychological symptoms, and academic and social integration in undergraduate students. *Journal of American College Health*, 64(5), pp. 362-370. <http://doi.org/10.1080/07448481.2016.1154559>
2. Alola, A. A., Alola, U. V. and Sarkodie, S. A. (2021). The COVID-19 and financial stress in the USA: Health is wealth. *Environment, Development and Sustainability*, 23(6), pp. 9367-9378. <https://doi.org/10.1007/s10668-020-01029-w>
3. American Psychological Association. (APA). (2022). *Stress in America 2022*. NE, Washington.
4. Britt, S. L., Canale, A., Fernatt, F., Stutz, K. and Tibbetts, R. (2015). Financial stress and financial counseling: Helping college students. *Journal of Financial Counseling and Planning*, 26(2), pp. 172-186. <https://doi.org/10.1891/1052-3073.26.2.172>
5. Bronfenbrenner, U. (1979). *The ecology of human development: Experiments by nature and*

- design*. Harvard University Press, London, England.
6. Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., ... and Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, 2(9), pp. 637-644. <https://doi.org/10.1038/s41562-018-0399-z>
 7. Cohen, S. and Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin*, 98(2), pp. 310-357. <https://doi.org/10.1037/0033-2909.98.2.310>
 8. Conger, R. D., Elder, G. H., Jr., Lorenz, F. O., Conger, K. J., Simons, R. L., Whitbeck, L. B., Huck, S., and Melby, J. N. (1990). Linking economic hardship to marital quality and instability. *Journal of Marriage and Family*, 52(3), pp. 643-656. <https://doi.org/10.2307/352931>
 9. Das, D., Kumar, S. B., Tiwari, A. K., Shahbaz, M. and Hasim, H. M. (2018). On the relationship of gold, crude oil, stocks with financial stress: A causality-in-quantiles approach. *Finance Research Letters*, 27, pp. 169-174. <https://doi.org/10.1016/j.frl.2018.02.030>
 10. French, D. and McKillop, D. (2017). The impact of debt and financial stress on health in Northern Irish households. *Journal of European Social Policy*, 27(5), pp. 458-473. <https://doi.org/10.1177/0958928717717657>
 11. Gelman, A. and Carlin, J. (2014). Beyond power calculations: Assessing type S (sign) and type M (magnitude) errors. *Perspectives on Psychological Science*, 9(6), pp. 641-651. <https://doi.org/10.1177/1745691614551642>
 12. Grosz, M. P., Rohrer, J. M. and Thoemmes, F. (2020). The taboo against explicit causal inference in nonexperimental psychology. *Perspectives on Psychological Science*, 15(5), pp. 1243-1255. <https://doi.org/10.1177/1745691620921521>
 13. Haghi, M. (2026). Teacher self-efficacy and teacher burnout in ESP higher education contexts: A systematic. *Journal of Cognition, Emotion, & Education*, 4(1), pp. 52-63. <https://doi.org/10.22034/cee.2026.578681.1043>
 14. Hobfoll, S. E. (2001). The influence of culture, community, and the nested-self in the stress process: Advancing conservation of resources theory. *Applied Psychology*, 50(3), pp. 337-421. <https://doi.org/10.1111/1464-0597.00062>
 15. Huang, R., Ghose, B. and Tang, S. (2020). Effect of financial stress on self-rereported health and quality of life among older adults in five developing countries: A cross sectional analysis of WHO-SAGE survey. *BMC Geriatrics*, 20(1), A. 288. <https://doi.org/10.1186/s12877-020-01687-5>
 16. Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), A. e124. <https://doi.org/10.1371/journal.pmed.0020124>
 17. Kim, J., Garman, E. T. and Sorhaindo, B. (2003). Relationships among credit counseling clients' financial wellbeing, financial behaviors, financial stressor events, and health. *Journal of Financial Counseling and Planning*, 14(2), pp. 75-87.
 18. Liu, Y. and Merritt, D. H. (2018). Familial financial stress and child internalizing behaviors: The roles of caregivers' maltreating behaviors and social services. *Child Abuse & Neglect*, 86, pp. 324-335. <https://doi.org/10.1016/j.chiabu.2018.09.002>
 19. Mahmudah, F. N., Putra, E. C. and Wardana, B. H. (2021). The impacts of COVID-19 pandemic: External shock of disruption education and financial stress cohesion. *FWU Journal of Social Sciences*, 15(2), pp. 42-64. <http://doi.org/10.51709/19951272/Summer-2/3>
 20. Mensi, W., Rehman, M. U. and Vo, X. V. (2022). Impacts of COVID-19 outbreak, macroeconomic and financial stress factors on price spillovers among green bonds. *International Review of Financial Analysis*, 81, A. 102125. <https://doi.org/10.1016/j.irfa.2022.102125>
 21. Mikkonen, J. and Raphael, D. (2010). The Canadian facts. *Toronto, ON: York University School*

- of Health Policy and Management*, 1(1), pp. 1-63.
22. Montpetit, M. A., Kapp, A. E. and Bergeman, C. S. (2015). Financial stress, neighborhood stress, and well-being: Mediation and moderation models. *Journal of Community Psychology*, 43(3), pp. 364-376. <https://doi.org/10.1002/jcop.21684>
 23. Mullainathan, S. and Shafir, E. (2013). *Scarcity: Why having too little means so much*. Macmillan, London, United Kingdom.
 24. Nasr, R., Rahman, A. A., Haddad, C., Nasr, N., Karam, J., Hayek, J., ... and Alami, N. (2024). The impact of financial stress on student wellbeing in Lebanese higher education. *BMC Public Health*, 24(1), A. 1809. <https://doi.org/10.1186/s12889-024-19312-0>
 25. Page, M. J., McKenzie, J. E., Bossuyt, P. M., et al. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, A. 71. <http://dx.doi.org/10.1136/bmj.n71>
 26. Patel, M. R., Kruger, D. J., Cupal, S. and Zimmerman, M. A. (2016). Effect of financial stress and positive financial behaviors on cost-related nonadherence to health regimens among adults in a community-based setting. *Preventing Chronic Disease*, 13, A. E46. <http://dx.doi.org/10.5888/pcd13.160005>
 27. Pishghadam, R. (2025). Magnitude thinking, identity, and language: Reassessing perceptions in migration contexts. *International Journal of Society, Culture & Language*, 13(1), pp. 1-10. <https://doi.org/10.22034/ijsc.2025.721091>
 28. Pishghadam, R. (2026). Introducing magnitude research: A transformative framework for business studies. *Journal of Business, Communication and Technology*, 5(1), pp. 1-12. <https://doi.org/10.56632/bct.2026.5101>.
 29. Pishghadam, R. and Ebrahimi, S. (2026). Magnitude psychology: A cognitive framework for proportional judgment in reflective education. *Journal of Cognition, Emotion & Education*, 4(1), pp. 1-20. <https://doi.org/10.22034/cee.2026.579070.1044>
 30. Richardson, T., Elliott, P., Roberts, R. and Jansen, M. (2017). A longitudinal study of financial difficulties and mental health in a national sample of British undergraduate students. *Community Mental Health Journal*, 53(3), pp. 344-352. <https://doi.org/10.1007/s10597-016-0052-0>
 31. Robb, C. A. (2017). College student financial stress: Are the kids alright?. *Journal of Family and Economic Issues*, 38(4), pp. 514-527. <https://doi.org/10.1007/s10834-017-9527-6>
 32. Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, 1(1), pp. 27-42. <https://doi.org/10.1177/2515245917745629>
 33. Rokhsari, S. (2025). A study of GCC economic visions through magnitude thinking. *Journal of Business, Communication and Technology*, 4(2), pp. 18-34. <https://doi.org/10.56632/bct.2025.4202>
 34. Simmons, J. P., Nelson, L. D. and Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), pp. 1359-1366. <https://doi.org/10.1177/0956797611417632>
 35. Simonse, O., Van Dijk, W. W., Van Dillen, L. F. and Van Dijk, E. (2024). Economic predictors of the subjective experience of financial stress. *Journal of Behavioral and Experimental Finance*, 42, A. 100933. <https://doi.org/10.1016/j.jbef.2024.100933>.
 36. Smathers, K., Chapman, E., Deringer, N. and Grieb, T. (2024). The relationship between financial stress and college retention rates. *Journal of College Student Retention: Research, Theory & Practice*, 26(2), pp. 581-604. <https://doi.org/10.1177/15210251221104984>
 37. Swigonski, N. L., James, B., Wynns, W. and Casavan, K. (2021). Physical, mental, and financial stress impacts of COVID-19 on early childhood educators. *Early Childhood Education Journal*,

- 49(5), pp. 799-806. <https://doi.org/10.1007/s10643-021-01223-z>
38. Thayer, Z. M. and Gildner, T. E. (2021). COVID-19-related financial stress associated with higher likelihood of depression among pregnant women living in the United States. *American Journal of Human Biology*, 33(3), A. e23508. <https://doi.org/10.1002/ajhb.23508>
39. Van Dijk, W. W., van der Werf, M. M. and van Dillen, L. F. (2022). The psychological inventory of financial scarcity (PIFS): A psychometric evaluation. *Journal of Behavioral and Experimental Economics*, 101, A. 101939. <https://doi.org/10.1016/j.socec.2022.101939>
40. Whitcomb, D., Battaly, H., Baehr, J. and Howard-Snyder, D. (2017). Intellectual humility: Owning our limitations. *Philosophy and Phenomenological Research*, 94(3), pp. 509-539. <https://doi.org/10.1111/phpr.12228>
41. White Jr, K. J. (2020). Financial stress and the relative income hypothesis among Black college students. *Contemporary Family Therapy*, 42(1), pp. 25-32. <https://doi.org/10.1007/s10591-019-09531-8>
42. Wilson, S., Hastings, C., Morris, A., Ramia, G. and Mitchell, E. (2023). International students on the edge: The precarious impacts of financial stress. *Journal of Sociology*, 59(4), pp. 952-974. <https://doi.org/10.1177/14407833221084756>
43. Xiao, J. J. and Kim, K. T. (2022). The able worry more? Debt delinquency, financial capability, and financial stress. *Journal of Family and Economic Issues*, 43(1), pp. 138-152. <https://doi.org/10.1007/s10834-021-09767-3>