Ferdowsi University of Mashhad
Press (FUMP)

IJAAF

IRANIAN JOURNAL OF

ACCOUNTING, AUDITING and FINANNCE

Volume 10, Issue 2, Spring 2026

Serial Number: 35

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

Title	Authors	Page Number
Bibliometric Analysis of Fractal Patterns in Stock Markets: Trends and Perspectives	**Mohammad Javad Nourahmadi Marziyeh Nourahmadi**	**1-20**
The Impact of the COVID-19 Crisis on Corporate Performance: The Moderating Role of Cash and Human Resources	**Najmeh Hajian Mahdieh Yazdanbakhsh**	**21-36**
The Analysis of Peer Effects in Environmental, Social, and Governance Disclosure: The Moderating Role of Firm Experience and Size	**Hamed Parvizikia Daruosh Foroghi Narges Hamidian Hasan Fattahi Nafchi**	**37-54**
The Impact of Conspiracy Illusion on Auditor Independence: The Moderating Roles of Tolerance for Ambiguity and Religious Orientation	**Shirin Molavi Ameneh Bazrafshan**	**55-67**
A Comparative Study of XGBoost and Artificial Neural Networks for Earnings Management Prediction	**Maedeh Moayeri Mohammad Arabmazar Yazdi Vahid Menati**	**69-87**
A Deep Learning Framework to Model the Moderating Effect of Ethical Leadership on Emerging Technology's Impact on Auditors' Professional Judgment	**Amir Avazpour Hashem Valipour Hamid Salehi**	**89-110**
The Impact of The Revolving Door Phenomenon on the Financial Performance	**Hadi Shafii Dizaji Gholamreza Mansourfar**	**111-124**
Predicting Stock Market Returns Using Temporal Fusion Transformer: A Comprehensive Data-Driven Approach	**Somayeh Mohebi Mohammad Reza Mohebbi Javad Mohebbi Najm Abad**	**125-151**
The Effect of Working Capital Management on Cost Stickiness Considering the Mediating Role of Financial Constraints	**Mehdi Kazempour Barough Hamzeh Didar Farzad Ghayour**	**153-170**
Audit Fees under Dual Earnings Management: The Moderating Role of Accruals in the REM-Audit Fee Nexus	**Seyed Hamid Jalali Hassan Hemmati Omid Faraji Hossein Jabbari**	**171-195**

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 2, Spring 2026, Serial 35

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
Bibliometric Analysis of Fractal Patterns in Stock Markets: Trends and Perspectives	Mohammad Javad Nourahmadi Marziyeh Nourahmadi	1
The Impact of the COVID-19 Crisis on Corporate Performance: The Moderating Role of Cash and Human Resources	Najmeh Hajian Mahdieh Yazdanbakhsh	21
The Analysis of Peer Effects in Environmental, Social, and Governance Disclosure: The Moderating Role of Firm Experience and Size	Hamed Parvizikia Daruosh Foroghi Narges Hamidian Hasan Fattahi Nafchi	37
The Impact of Conspiracy Illusion on Auditor Independence: The Moderating Roles of Tolerance for Ambiguity and Religious Orientation	Shirin Molavi Ameneh Bazrafshan	55
A Comparative Study of XGBoost and Artificial Neural Networks for Earnings Management Prediction	Maedeh Moayeri Mohammad Arabmazar Yazdi Vahid Menati	69
A Deep Learning Framework to Model the Moderating Effect of Ethical Leadership on Emerging Technology's Impact on Auditors' Professional Judgment	Amir Avazpour Hashem Valipour Hamid Salehi	89
The Impact of The Revolving Door Phenomenon on the Financial Performance	Hadi Shafii Dizaji Gholamreza Mansourfar	111
Predicting Stock Market Returns Using Temporal Fusion Transformer: A Comprehensive Data-Driven Approach	Somayeh Mohebi Mohammad Reza Mohebbi Javad Mohebbi Najm Abad	125
The Effect of Working Capital Management on Cost Stickiness Considering the Mediating Role of Financial Constraints	Mehdi Kazempour Barough Hamzeh Didar Farzad Ghayour	153
Audit Fees under Dual Earnings Management: The Moderating Role of Accruals in the REM-Audit Fee Nexus	Seyed Hamid Jalali Hassan Hemmati Omid Faraji Hossein Jabbari	171

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 Staszkiewicz
Professor of Accounting, Warsaw, Poland

Kayhan Tajeddini
Professor of Strategic Mangement, 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 writhen 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 Roman11and 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

IJAAF; Vol. 10, No. 2, Spring 2026, Serial 35

Ferdowsi University of Mashhad

Iranian Journal of
Accounting, Auditing & Finance

Quarterly

RESEARCH ARTICLE

Bibliometric Analysis of Fractal Patterns in Stock Markets: Trends and Perspectives

Mohammad Javad Nourahmadi*
Department of Theoretical Economics, Faculty of Economics, Allameh Tabatabaii University, Tehran, Iran
Marziyeh Nourahmadi
Department of Financial Engineering, Hazrat-e Masoumeh University, Qom, Iran

How to cite this article:
Nourahmadi, M. J. and Nourahmadi, M. (2026). Bibliometric Analysis of Fractal Patterns in Stock Markets: Trends and Perspectives. Iranian Journal of Accounting, Auditing and Finance, 10(2), 1-20. doi: 10.22067/ijaaf.2026.47633.1590
https://ijaaf.um.ac.ir/article_47633.html

ARTICLE INFO

Article History
Received: 2025-03-31
Accepted: 2025-09-15
Published online: 2026-04-12

Abstract

Traditional financial models fail to capture the nonlinear, self-similar patterns in stock markets, whereas the Fractal Market Hypothesis (FMH) provides a more robust framework. This study aimed to carry out a comprehensive bibliometric analysis to map the intellectual structure and evolution of research on stock markets with fractal patterns. Based on 1,280 documents obtained from the Scopus database (1982–2025), bibliometric analyses were conducted, including co-authorship and co-word analysis, as well as thematic evolution with the Bibliometrix R package. Despite limitations in study design, results indicated a significant shift from basic to their applied interdisciplinary research. The multifractal approach has emerged as the leading methodology with around 42% studies after 2005. Research output is led by China, India and the USA with increasing interest in financial crises, cryptocurrencies and hybridizing fractals with artificial intelligence. The bibliometric results yield empirical evidence that reinforces the applicability of the FMH in addressing market complexity and aligning with contemporary financial phenomena. Analysis of common themes in literature highlights certain knowledge deficiencies, predominantly regarding comparative studies between developed and emerging markets, dynamic performance under systemic shocks of fractal structures and the utility function analysis based on where fractal-based trading strategies are implemented. This is clearly a stepping stone for more extensive research on that topic, but it also highlights how fractal approaches can span across academic fields in finance.

Keywords:
Bibliometric Analysis, Fractal Patterns, Multifractality, Market Efficiency, Stock Market

https://doi.org/10.22067/ijaaf.2026.47633.1590

CC BY

NUMBER OF REFERENCES
42

NUMBER OF FIGURES
9

NUMBER OF TABLES
1

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

*Corresponding Author:
Mohammad Javad Nourahmadi
Email: mjnourahmadi@atu.ac.ir
Tel: 09155164202
ORCID: 0000-0002-0738-9069

1. Introduction

Financial markets are examples of complex adaptive systems with nonlinear dynamics, feedback processes, and changing interdependencies. Sector luminaries such as Fama (1970) rendered textbook paradigms, including the Efficient Market Hypothesis (EMH), which state that prices instantly and completely reflect all available information and asset returns are essentially random. Although the EMH presents a foundational understanding of contemporary finance, the underlying assumptions regarding perfect rationality and equilibrium-based efficiency cannot account for these anomalies or inefficient patterns, volatility clustering, and long-memory effects (Onali and Goddard, 2011; Brătian et al., 2021). The repetitive occurrence of crises, speculative bubbles and asymmetric information flows only serves to reinforce the notion that market behavior does indeed differ from the EMH theoretical ideal (Karp and Van Vuuren, 2019).

Given these empirical failings, the Fractal Market Hypothesis (FMH) did not just represent an alternative, but a necessarily progressive development of financial ideas. Drawing from Mandelbrot's (1972) idea of self-similarity and Peters' (1989) initial fractal approach to the structure of market data, FMH redefines financial markets as self-organized structures governed by investors with heterogeneous time horizons. Whereas EMH assumes informational homogeneity, FMH allows for the insights that stability and liquidity require a balance between long- and short-term investment perspectives. This balance gets disrupted, as in the crisis markets display a multifractal scaling nature that elicits nonlinear responses and extreme volatility (Berghorn et al., 2019; Okorie and Lin, 2021). Therefore, FMH delivers a more accurate explanation by combining heterogeneity, memory and dynamic adjustment found in data that EMH has systematically neglected.

In the past four decades, the collaborative effort of multifractal analysis, econophysics and computational finance has been iterative to form FMH (Jiang et al., 2019; Zhou et al., 2024), pushing further theoretical study and empirical modeling from regime paradigm to complexity in a market. Further modern analysis also displays that FMH identifies structural shifts in markets affected by global shocks like the COVID-19 pandemic and new developing phenomena in digital assets and decentralized finance (Okorie and Lin, 2021; ElNabulsi and Anukool, 2025). Additionally, the relationship of fractal modeling with artificial intelligence and quantum computation renders FMH as a comprehensive approach for exploring financial systems operating under uncertain conditions where nonlinear interactions take place (Popovska Georgieva-Tsaneva 2024; Kehinde et al., 2023).

Despite these developments, some research gaps remain. Existing studies comparing the performance of emerging and developed markets are limited, as is the dynamicity behind the multi-fractal structures' response to systemic shocks. Moreover, a potential area for future research is the marriage of FMH-based indicators with machine learning and big data analytics methodologies to achieve real-time forecasting and risk management (Leventides et al., 2024). To tackle these problems, we need a broad understanding of how the FMH has developed conceptually and empirically over time.

The current study fills these gaps by using a bibliometric method to explore the intellectual structure and evolution of research on fractal stock markets between 1982 and 2025. Bibliometric analysis has been acknowledged as a qualitative and quantitative approach that can identify multiple research trends, conceptual clusters and knowledge gaps by examining scientific publication data such as articles, citations through papers' keywords, and collaboration networks (Nourahmadi and Rasti, 2025; Nourahmadi, 2024). Specifically, it investigates the following four research questions:

1. What are the temporal trends in FMH-related publications, and how do thematic pathways evolve across different periods?

2. Which authors, countries, and journals have the strongest influence based on co-authorship networks, citation counts, and collaboration patterns?
3. What conceptual clusters can be identified through keyword co-occurrence and thematic mapping, and how do they relate to emerging research fronts?
4. What are the knowledge gaps based on bibliometric evidence and what suggestions for future directions could inform interdisciplinary and applied research in FMH?

This study addresses these questions and helps consolidate existing knowledge, while identifying promising ways forward for interdisciplinary and applied research.

2. Theoretical framework

2.1. Origins of fractal geometry and its introduction to financial economics

In 1975, Benoit Mandelbrot introduced the amazing mathematical construction of fractals, coining a new term (now it does). He was also remarkably interested in using mathematical language to describe unusual patterns found at various scales in nature, such as coastlines, clouds and mountains whose apparent complexities resist Euclidean geometry. He introduced the idea of fractals, shapes that exhibit similar features at various scales, a property called self-similarity. Now this idea was used not only in natural science but also in social and economic spheres. In his early work, including a 1972 paper, Mandelbrot demonstrated that cotton prices in financial markets exhibit fractal characteristics and that their fluctuations follow non-Gaussian distributions (Mandelbrot, 1972). Fractal geometry fundamentally changed financial modeling by providing a quantitative link between market irregularities and scaling behavior, enabling researchers to capture nonlinear dynamics and volatility clustering**.**

The late 1980s saw an acceleration in the adoption of fractal geometry as a way to characterize asset price dynamics, notably by the financial economist Edgar Peters. In particular, Peters, in a seminal 1989 article in the Financial Analysts Journal, showed how market price data exhibited fractal structures. Unlike the classical models, like Geometric Brownian motion, financial markets show nonlinear, complex behavior that is better described using fractal instruments (Peters 1989). As a result, Fractal Market Hypothesis (FMH) was created, being proposed as an alternative hypothesis of the Efficient Market Hypothesis (EMH). FMH requires that markets are not perfectly efficient and supports the existence of self-similarity and long memory in price behavior (Evertsz, 1995; Malik, 2011).

Within this framework, price changes should be examined across multiple time scales, as investors with different horizons, short-term (e.g., day traders), medium-term (e.g., institutional investors), and long-term (e.g., pension funds) influence the market. These complex interactions generate fractal patterns in which small daily fluctuations can repeat larger weekly or monthly patterns. This multi-scale behavior forms the basis for modeling volatility and forecasting market regimes.

2.2. Efficient market hypothesis versus fractal market hypothesis

The Efficient Market Hypothesis (EMH) claims that the price of any asset fully reflects all known information at any point in time, which was proposed by Eugene Fama in the 1970s. The EMH hypotheses that states that no investor should be able to achieve a systematic return greater than the market due to price movement being a random and therefore unpredictable process that can exist in a weak, semi-strong or strong form (Fama, 1970). In contrast, experience during the last thirty years has demonstrated some consistent structural inefficiencies in markets, e.g., calendar effects, price bubbles and financial crises. These inefficiencies show that markets always efficiently weigh all of the available information and investor irrationality as a major factor (Onali and Goddard, 2011).

By contrast, in FMH, price data exhibit the fractal structures that are financial markets being made up of investors with a diversity of time horizons (Peters, 1989). FMH espouses the long memory and self-similarity of market volatility— that is, the behavior of markets across different time intervals is similar, if not in intensity or magnitude. In FMH, financial markets operate under different regimes: when the market is in a normal state, stable market conditions prevail due to the stabilizing effect of diverging investor horizons, and crises are characterized by an (over)concentration on short-term investment goals resulting in high returns volatility (Berghorn et al., 2019). In contrast to EMH, the FMH incorporates resetting feedback loops, the memory effect and dynamics adaptation, resulting in a more general conceptual apparatus for modeling market imbalance phenomena (Karp and Van Vuuren 2019; Okorie and Lin 2021).

Empirical studies confirm FMH's superiority in explaining crisis dynamics and volatility clustering in emerging markets (ElNabulsi and Anukool, 2025).

2.3. Key concepts in fractal theory

In this section, we provide an overview of essential concepts that form the basis for the analysis of fractal markets to connect theoretical foundations with analytical tools.

Self-Similarity: The property of having repeating patterns across different time scales, which carries over to financial markets as well, where price behavior across days, weeks, or months is similar but the distribution of prices is different (Evertsz, 1995).

Long Memory: The Hurst exponent is used in measuring long memory. If $H > 0.5$, there is a tendency for trends to persist, and if $H < 0.5$, then mean reversion is expected (Yu et al., 2006). This is a key property in FMH's predictive power.

Fractal Dimension: measures how complex/rough a time series; higher numbers express more volatile and irregular market structures (Bayraktar et al., 2004). This metric allows for clearly distinguishing between market efficiency in various regions.

Richtigen: MFDFA and MFDCCA methods show asymmetries and nonlinear patterns (Jiang et al., 2019; Zhou et al., 2024). Multifractality is especially effective in detecting regime shifts and systemic risks.

2.4. Analytical frameworks

In this section, we present statistical models, which are based on fractal theory, that can be used in financial analysis.

ARFIMA Model: Autoregressive Fractionally Integrated Moving Average models are capable of modeling long memory processes typical in financial time series (Panas, 2001). They extend ARIMA to allow for fractional differencing, which makes them applicable where there are pronounced trends.

Wavelet Analysis: Extracts multi-scale features from complex time series, working great when dealing with high-frequency data, especially to analyze volatility and fractal dimension (Bayraktar et al., 2004).

MFDFA and MFDCCA: Multifractal Detrended Fluctuation Analysis and its cross-correlation variant have been applied extensively to discover multifractal structures as well as dependencies between assets (Zhou et al., 2024). The combination of these methods may augment the ARFIMA phenomenon and wavelet analysis to identify several new facts with a scale sensitive interests about market.

To improve clarity, Section 2.5 focuses on geometric and dynamic models, while Section 2.4 emphasizes statistical and time-series approaches.

2.5. Advanced models

This section introduces dynamic and geometric models that extend fractal analysis into new theoretical domains. Each model is presented using a consistent format.

Fractional Brownian Motion (FBM): Definition: A generalization of classical Brownian motion with long memory. Insights: Models time series and cross-sectional complex dependencies in financial data. Use case: Used to model stock indices more accurately than conventional approaches (Brătian et al., 2021; Gu, 2025).

Geometric Arbitrage Theory: Definition: A geometric structure that models arbitrage as curvature. Goal: Use differential geometry to illuminate market dynamics. Application: Used to identify arbitrage opportunities in cryptocurrency markets (Farinelli, 2015; ElNabulsi and Anukool, 2025).

Higher conceptual coherence and theoretical integration are achieved as this section ensures stylistic consistency among the various models and clarifies the purposes of each model.

2.6. Literature review

2.6.1. Classical studies

Peters (1989) initially introduced fractal structures in financial markets through the early studies. Drawing on Mandelbrot's work in fractal geometry, he showed that asset price variations are self-similar, a property that never quite gets captured with classical models. Then, Evertsz (1995) quantitatively highlighted the relevance of fractal dimensions in financial time series analysis. Huang and Yang (1995) further illustrated by examining multinational markets that structures of stock returns exhibit a multi-scale property. Such studies laid the groundwork that helped establish a new field of research, which over subsequent decades has grown to an increasingly frenetic pace. These works have been cited as foundational for the combination of fractal models with artificial intelligence (Blackledge and Lamphiere, 2021) in recent reviews. Yet these initial studies, although valuable, did not have strong empirical validation across many states of the market, limiting generalizability.

2.6.2. Developed vs. emerging markets

The distinction between developed and emerging markets has been a key focus of fractal research. Bianchi and Frezza (2017) showed that market efficiency is dynamic, and fractal characteristics can reveal the relative efficiency or inefficiency of markets. Samadder et al. (2013) concentrated on Indian stock market indices and discovered strong self-similarity of price series. More recently, Xu et al. In their research, (2022) found asymmetric fractal features for the style indices of China are present, indicating that market efficiency plays a significant role under certain conditions. These have shown that EMs tend to be more complex and multifractal, indicating structural inefficiencies and weak information flow. Evidence from the MFDFA of African markets affirmed these disparities and indicated to policymakers that they could also utilize these tools for improved efficiency (Moradi et al., 2021). However, comparative studies tend to be limited in both methodology and data granularity, so conclusions across markets are tentative.

2.6.3. Market efficiency and asymmetry

A wide range of fractal data has been used to test the Efficient Market Hypothesis (EMH). Onali and Goddard (2011) studied European markets, finding significant inefficiencies in many indices that can be detected using fractal techniques. Gayathri et al. (2013) investigated the Indian market and demonstrated that fractal dimensions can significantly explain non-linear market dynamics. Grönlund et al. (2012) developed a "Fractal Profit Landscape," which means that trading systems are developed in an environment determined by fractals (2012). Together, these studies have shown that the markets are not fully efficient, especially during times of volatility or crisis and fractal analysis offers a useful tool for detecting such inefficiencies. More recent work using this approach incorporates information entropy to highlight asymmetries in return distributions (Liu et al. 2025).

To improve thematic flow, the next section builds on these insights by examining how crises and systemic shocks further amplify fractal complexity.

2.6.4. Crises and economic shocks

A critical area of research concerns the impact of crises and systemic events on the fractal structure of markets. Okorie and Lin (2021) demonstrated that the COVID-19 pandemic significantly altered fractal features and increased market complexity. Moradi et al. (2021), who examined several crises for the Tehran and London markets, also suggested that the Fractal Market Hypothesis can help to predict stock returns under high-risk conditions. Yang (2025) took a crack at the "fractal DNA" viral concept to account for Albion pattern-matching crisis arcs evolving hierarchically in fractal time. In studying the effects of mergers and acquisitions on stock prices, Verma and Kumar (2023) analyzed changes in the fractal dimensions concerned with market complexity. Our analysis of the EU trade data provides shock studies and shows that its fractal structures fundamentally change under shocks to give rise to new topological order in both short memory and long memory. Analyzing the multifractality of South African sector indices revealed that higher multifractality was accompanied by crisis periods (Amini et al., 2024). However, studies specifically focusing on crises are frequently retrospective in nature, limiting their predictive power.

2.6.5. Practical applications and trading strategies

Fractal analysis is not only a theoretical concept. The study conducted by Blackledge and Lamphiere (2021) gives an extensive overview of the various applications in FMH for algorithmic trading, showing that fractal patterns can be used in price forecasting/management and risk management. Popovska and Georgieva-Tsaneva (2024) suggested robotic trading strategies considering fractional derivatives and volatility analysis. Gu (2025) proposed multi-timeframe fractal patterns for optimizing cross-asset trading. Liu et al. (2025) "Energy-Spacetime" architecture generalizes as fractal-game theory physics for markets. Recent research proposes to apply fractal analysis in cryptocurrency trading as a way of finding structure within chaotic price movements (Ko and Kuo-Shing, 2025; Kristjanpoller and Tabak, 2025). Additionally, combining fractals and filters based on Fractal Chaos Bands mitigates market noise and promotes multi-timeframe strategies (Blackledge and Lamphiere, 2021; Popovska and Georgieva-Tsaneva, 2024). While very promising, these methods have limited validation in real time and may struggle to adapt to regime shifts.

2.6.6. Market modeling and simulation

Researchers have sought to simulate markets using fractal-based models to better understand complex behavior. Leventides et al. (2024) simulated the Athens Stock Exchange using fractional Brownian motion. Amini et al. such as (2024) proposed a model explaining volatility clustering using fractals. A fresh approach using fractal-based qualitative financial modelling under high-risk scenarios was introduced in ElNabulsi and Anukool (2025). These approaches suggest that financial markets need to be represented not just as aggregates of trades, but rather as complex systems with behaviors across multiple scales. When modeling financial markets, Laplace–Mittag-Leffler distributions have proven useful (Kristjanpoller and Tabak, 2025). However, simulation models are typically never compromised due to parameter calibration and real-world comparability.

2.6.7. Complementary concepts and interdisciplinary approaches

Econophysics has also contributed to the advancement of fractal research. The field applies ideas from statistical physics and chaos theory to better understand how financial systems behave. Formally, financial markets tend to display complex behaviors similar to those of physical systems, such as volatility clustering, fat-tails and non-Gaussianity. Chaos Theory stresses great sensitivity to starting conditions, meaning that marginal variations in input can yield dramatically different results. That view is aligned with fractal geometry, which explains the nonlinearity and complex behavior (Castilhos et al., 2017). In 2025, we applied econophysics based on fractal analysis to African stock markets, as they are stated to have lower efficiency.

Farinelli (2015) used the tools from differential geometry to define arbitrage as the "curvature" of a principal fibre bundle and called this theory his Geometric Arbitrage Theory. This geometrical viewpoint complements the concept of fractals because both seek to uncover hidden order in financial data. Combining these theories can create all-new moulds in the market of stability analysis and arbitrage opportunity finding.

Together, these interdisciplinary frameworks—FMH, chaos theory, econophysics, and geometric modeling—form a conceptual ecosystem for understanding financial complexity. Their integration under the umbrella of fractal analysis enables multi-scale modeling, cross-asset forecasting, and systemic risk detection.

2.6.8. Cryptocurrencies and decentralized finance

In addition, the introduction of cryptocurrencies and blockchain technology has created new paths for fractal research. Considering the extreme volatility and nonlinear behavior of cryptocurrencies, a fractal approach may be sufficient (Jiang et al., 2019). Recent research has suggested that Bitcoin and other cryptocurrencies share multifractal character with long memory, similar to what we see in the traditionalization of emerging market movements. Moreover, with the recent injection of decentralized finance (DeFi), fully algorithm- and smart-contract-driven markets are emerging. By understanding the fractal characteristics of these markets, it is possible to develop new tools for risk management and predictive modelling. Some others have tried to use fractal analysis in cryptocurrency trading, initiating discoveries of structural patterns contained in the price algorithm at a very volatile scale (Ko and Kuo-Shing, 2025; Kristjanpoller and Tabak, 2025).

2.6.9. Methodological challenges and limitations

Despite considerable progress, the use of fractal analysis in finance still has its issues:

- **Data Noise and Nonstationarity**: This is particularly true for financial time series, which are known to be noisy and nonstationary, potentially biasing any estimation of a fractal index like the Hurst exponent (Goyal, 2024).
- **Computational Complexity**: Techniques like MFDFA necessitate significant computational capacities, restricting scalability (Kehinde et al., 2023).
- **Interpretability**: Economic interpretation of fractal indicators does not come so clearly, especially

3. Data and methodology

The Bibliometrix method analyzes various patterns and models in bibliometrics and its applications to study scientific trends, research developments in collaboration with several others (Nourahmadi and Parsi 2025); it also allows scholars and authors to discover high-demand research areas and plan their research aligned with the most up-to-date scientific trends. Using bibliometric analysis, this study systematically analyzes research on fractal patterns in stock markets. Bibliometric analysis is a strong method to map the intellectual structures, identify the key authors and institutions, and trace evolutionary trends in a specific research domain (Donthu et al., 2021). This technique offers extensive descriptive as well as visual insights into the existing literature in this area, especially considering the very small number of publications on fractal finance which have emerged in recent decades.

Methods: The Scopus database, a well-known source for indexing scientific articles, was used to collect relevant scientific data. Data collection occurred in multiple stages:

Keyword Selection: A variety of both theoretical and applied keywords relevant to the research domain of interest were used to provide comprehensive coverage. The main combinations were:

- Fractal Pattern + Fractal Finance
- Fractal Pattern + Self-similarity
- Financial Market + Fractal Market
- Fractal Finance + Multifractal
- Multifractal Analysis + Market Volatility

These combinations span both theoretical dimensions (e.g., self-similarity, fractal dimension) and applied dimensions (i.e., market volatility and option pricing).

3.1. The exact Boolean search query used in Scopus was

("Fractal Pattern" and "Fractal Finance") or ("Fractal Pattern" and "Self-similarity") or ("Financial Market" and "Fractal Market") or ("Fractal Finance" and "Multifractal") or ("Multifractal Analysis" and "Market Volatility").

The data were extracted from Scopus on August 25, 2025.

Initial Data Retrieval: The initial search in Scopus retrieved 1,704 scientific documents, including journal articles, conference papers, book chapters, and reviews.

Data Filtering: The initial steps for ensuring quality and focusing on reliable scientific literature were as follows:

1. Removing duplicate documents resulting from keyword overlaps.
2. Applying a publication window from 1982 to 2025.
3. Excluding irrelevant documents based on title and keyword review.
4. Retaining only articles written in English and excluding publications in other languages.

The final dataset, after applying filtering measures, included 1,280 documents from 569 credible sources; the various types of articles included journal articles, conference papers, book chapters and reviews.

- Data Preparation for Analysis: Table Z-filtered data were exported into BibTeX and CSV formats to guarantee compatibility with bibliometric analysis tools. These comprised authors and co-authorship networks, year of publication, source journals, keywords, abstracts, citation counts and references. This data was used as the basis for descriptive statistics analysis, as well as for network-based analyses such as networks of keyword co-occurrence, thematic maps and temporal trend analyses.
- Analytical approach: We used the Bibliometrix package in R (Aria and Cuccurullo, 2017) for data analysis. Author Co-citation analysis, Collaboration Analysis and Country/Institutional Publication output were performed using the Bibliometrix package in RStudio, whereas its web-based interface, Biblioshiny, was used for interactive visualization as well as analysis. The employed techniques included:
- Descriptive Statistics: Annual growth, average citations and document types, which gives a broad overview of up-to-date trends of publications and impact, and the history of resource

development.

- Co-authorship Analysis: Exploring collaboration trends among authors and nations. This also shows the structure, density, and internationality of research networks in this field.
- Co-word Analysis: Conceptual structure mining in keyword co-occurrence networks. This method reveals co-occurring terms and emerging research fronts, along with intellectual connections.
- Thematic Mapping: Grouping topics into push, foundational, niche and emergent/declining themes. This aids the emergent knowledge domains in identifying their focus areas from the periphery or depreciating topics, thus providing an insight into a research area positioning strategy.
- Tune Released: Histories of the article in different time points, which further illustrate changing thematic priorities over time and highlight newly emerging paradigms or methodological fashions.

4. Findings

Table 1 presents descriptive statistics of the conducted research:

Table 1. Descriptive statistics

Description	Results
General Information about the Data	
Time Span	1982–2025
Sources (Journals, Books, etc.)	569
Documents	1,252
Annual Growth Rate (%)	10.600
Average Document Age	10.7 years
Average Citations per Document	24.310
Content of Documents	
Keywords Plus	5,380
Author Keywords	2,967
Authors	
Total Number of Authors	2,466
Single-Author Papers	155
Author Collaboration	
Single-Author Documents	199
Average Number of Co-authors per Document	2.870
International Collaboration (%)	19.65%
Document Types	
Journal Article	919
Article-Book	1
Article-Book Chapter	1
Article-Conference Paper	1
Book	9
Book Chapter	25
Conference Paper	268
Conference Paper-Article	3
Review	21

The bibliometric dataset comprised 1,252 publications released between 1982 and 2025 from 569 sources. The annual publication growth rate was 10.6%, and the average citations per document amounted to 24.3, reflecting the significant and sustained impact of research in the field of fractal finance. The majority document type was journal articles (919), while only 268 papers were conference documents, which play a complementary role in new topics. These statistics show the persistence of the discipline in terms of scientific growth and significant influence on financial markets using fractals, which should serve as a firm base to further analysis (literature) trends.

4.1. Results and discussion

This section provides an in-depth examination of the bibliometric data, taking each of the four research questions (RQs) stated within the Introduction in turn. Using the Bibliometrix R package, the results, based on 1,280 documents indexed in the Scopus database from 1982 to 2025, show the intellectual structure and some of their influential contributors as well as their thematic evolution in stock markets through fractal patterns. Discussing these findings, they are not only exhaustive results to be interpreted descriptively but rather an empirical material that highlights the increasing relevance of FMH as a competing robust paradigm for financial modelling.

RQ1: What are the temporal trends in FMH-related publications, and how do thematic pathways evolve across different periods?

The aggregate analysis of annual scientific production (Figure 1) demonstrates a constant trend of increase directly, answering RQ1 about the temporal trends. This path is quantitative and qualitative in terms of degenerating traditional paradigms such as Efficient Market Hypothesis (EMH) (Fama, 1970) into the Fractal Market Hypothesis (FMH) (Peters, 1989). Research productivity has been characterized by three separate phases:

- Formation Phase (1982–2000): This was a slow growth period, accounting for fewer than 20 articles annually. It concentrated on heuristic theories from Mandelbrot (1972) and Peters (1989), laying the foundation of ideas such as self-similarity and fractal dimension.
- Expansion Phase (2001–2015): Rapid growth occurred during this period, peaking at around 60 articles published per year, which correlates with the massive uptake of computational tools such as multifractal detrended fluctuation analysis (MF-DFA).
- Maturity and Diversification Phase (2016–2025): The number of annual publications reaches more than 75 per year, corresponding to a cumulative propagation ratio of 10.6% (Table 1). The average citation per document during this phase reached 24.31, indicating the high impact and longevity of the research.

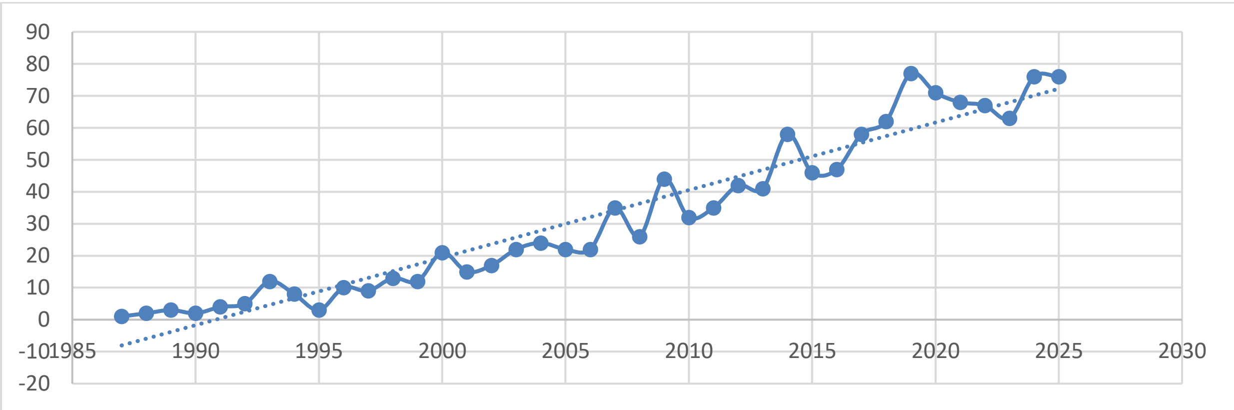

Figure 1. Annual Scientific Production

The three-phase analysis (Figure 1) shows the macro-level growth in publication volume, while a co-occurrence evolution map of themes over time (Figure 2) provides a more zoomed-in view of conceptual shifts that led to this growth. Collectively, they provide a panorama of the field's evolution. The map depicts five specific periods:

1. Period 1 (1982–2007): Focused on foundational mathematical models applied in basic sciences.
2. Period 2 (2008–2014): Marked the pivot toward financial economics, with core concepts like

"fractals" and "self-similarity" gaining prominence.

3. Period 3 (2015–2018): Defined by methodological diversification, integrating wavelet analysis and pattern recognition.
4. Period 4 (2019–2022): Consolidated research by applying advanced tools like detrended cross-correlation analysis (DCCA) to financial markets, particularly during crises like COVID-19.
5. Period 5 (2023–2025): Showcases a convergence with cutting-edge technologies, integrating fractal analysis with concepts from modern physics, artificial intelligence, and decentralized finance (DeFi).

The new themes of FMH from 2018 onwards, finally leading us to the merging with AI and DeFi, are evidence of how adaptable and relevant the FMH can be. This indicates that FMH is not a fixed theory, but a dynamic framework that can accommodate new methodologies to account for increasingly elaborate financial phenomena as an essential feature of theories with their base in complexity science.

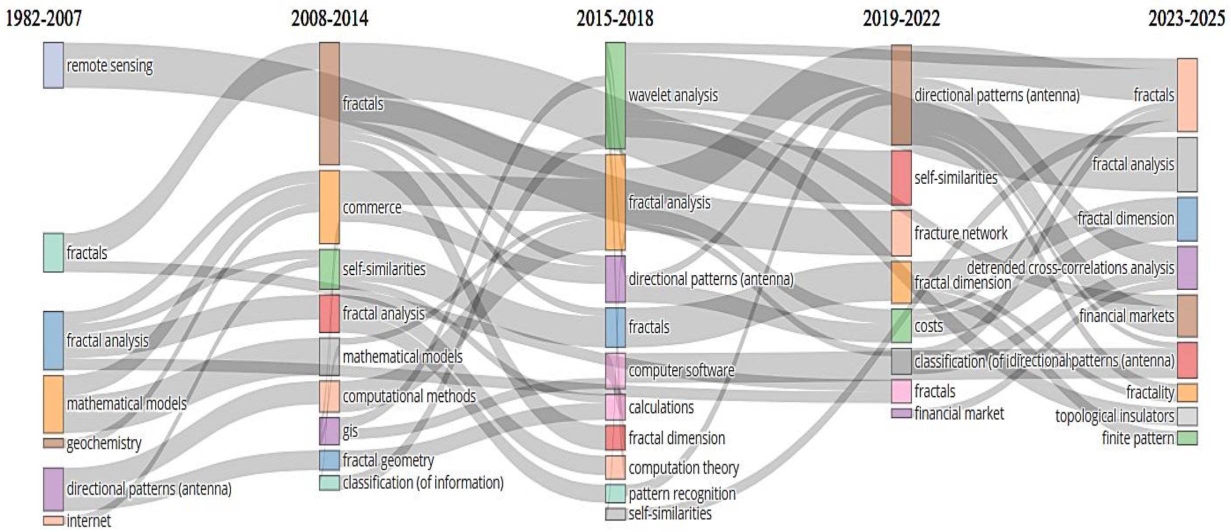

Figure 2. Thematic Evolution

RQ2: Which authors, countries, and journals have the strongest influence based on co-authorship networks, citation counts, and collaboration patterns?

The influence of authors and countries is mapped through the three-field plot (Figure 3) and the corresponding author's country chart (Figure 4). The analysis reveals a pronounced geographic concentration of knowledge production. China is the undisputed leader, with authors like Wang J, Zhou WX, and Jiang ZQ contributing a high volume of work. This prominence can likely be explained by large national investments into data science and econophysics, along with the high-profile teams at universities, and a tactical focus on complex systems. Poland has also shown high contributions (e.g., Drozdz S), whilst Brazil, following China, ranked as a top contributor (e.g., Fernandes LHS).

However, collaboration patterns vary considerably. Where China excels in single-country publications (SCP), countries such as Germany, the UK and Canada show particular willingness to work with other countries (MCP), with more than 47% from these nations co-authored internationally. This suggests that there is some bifurcation of research culture one the one hand, high-volume domestic output (China, India) and on the other hand, building international collaborative networks. This difference significantly affects how knowledge spreads globally and the potential for diverse viewpoints within the field.

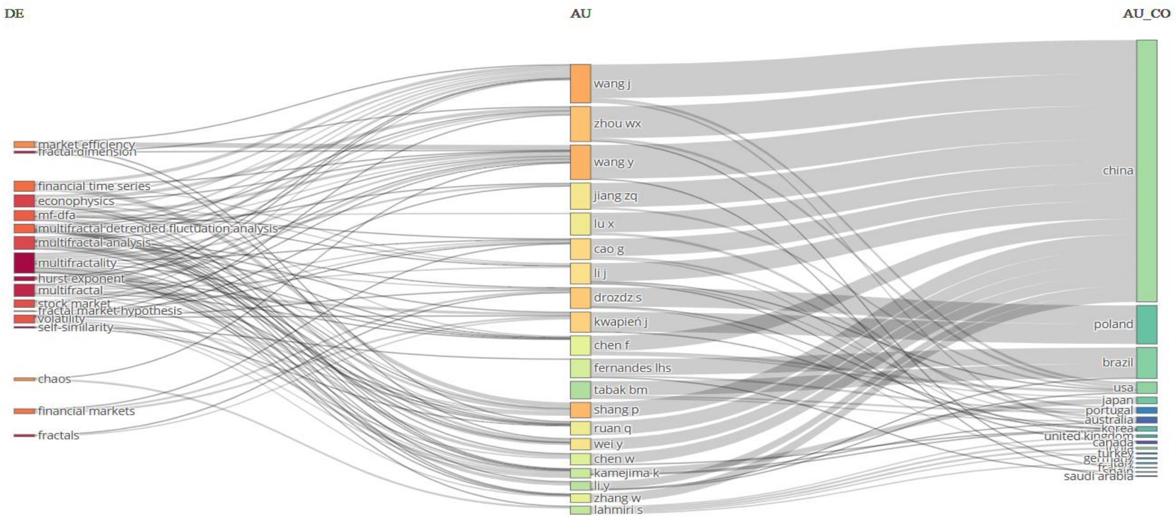

Figure 3. Three field plots

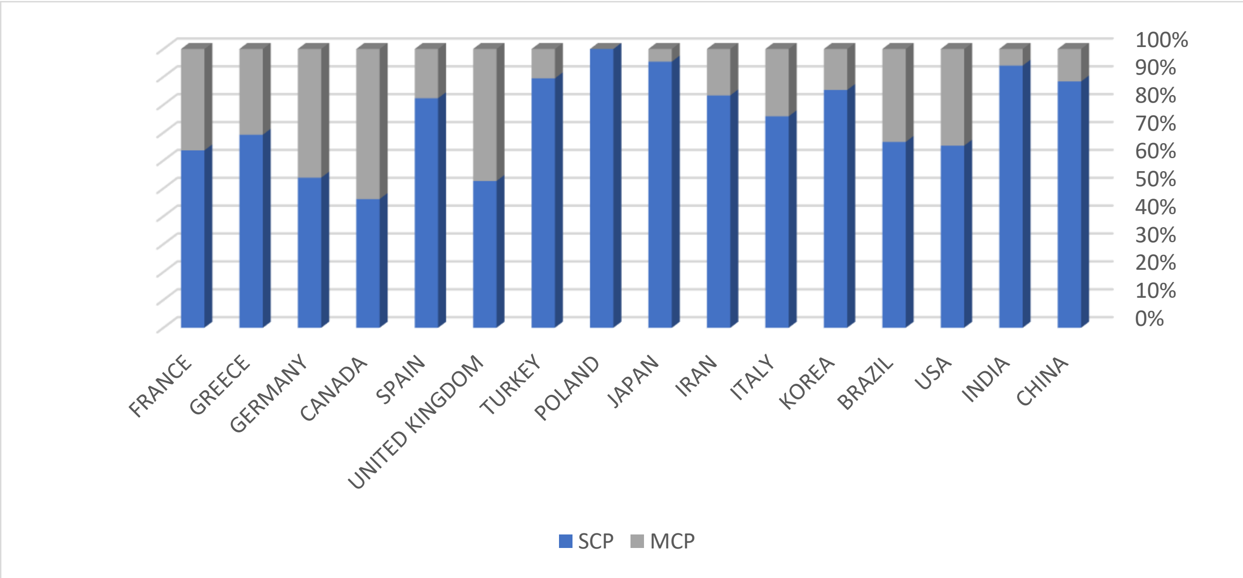

Figure 4. Corresponding author's countries

RQ3: What conceptual clusters can be identified through keyword co-occurrence and thematic mapping, and how do they relate to emerging research fronts?

Keyword analysis (Figures 5, 6, 7, 8, and 9) uncovers a rich conceptual structure with evolved significantly over time. The word cloud (Figure 5) and the trend topics (Figure 6) show a shift in focus. Specifically, the emphasis has moved away from core concepts like "self-similarity" and "fractal dimension" toward more practical areas such as "cryptocurrencies," "forecasting," and "risk management." The prominence of 'cryptocurrencies' and 'deep learning' alongside traditional terms signifies a critical paradigm shift in which the field is moving from pure financial econometrics towards an interdisciplinary convergence with data science, reflecting the need to analyze new, algorithm-driven markets. The emergence of terms like 'decentralized finance (DeFi)' in the most recent period illustrates the field's responsiveness to technological innovations and its expanding scope beyond traditional stock markets.

Figure 5. Word cloud

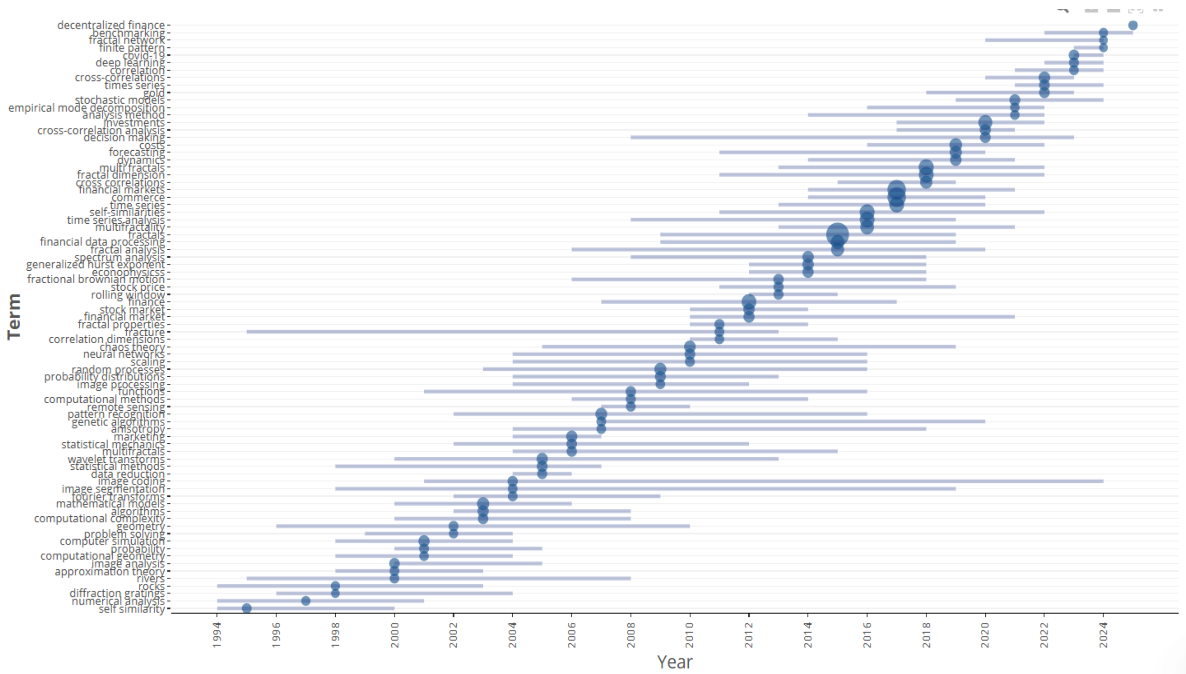

Figure 6. Trend topics

The co-occurrence network (Figure 7) organizes these keywords into four distinct clusters, each with distinct theoretical and practical relevance:

- Blue Cluster (Fundamental Concepts): This cluster represents the theoretical bedrock of the field, providing the fundamental language (e.g., self-similarity, chaos) that challenges the assumptions of traditional finance and underpins the FMH.

- Red Cluster (Applications in Financial Markets): Linking fractal tools directly with the core debates of finance, the Red Cluster is important because it provides empirical evidence contrary to the Efficient Market Hypothesis (EMH) and supportive towards a Fractal Market Hypothesis (FMH) framework (Onali and Goddard, 2011).
- Green Cluster (Multifractal Analyses◦Emerging Applications): The much-needed novelty, the applied evolution of the field, is suggested to be covered in a way offered by the Green Cluster. With a focus on cryptocurrencies and econophysics, it is essentially a research front seeking to model extreme volatility and systemic risk in contemporary, complex financial ecosystems.
- Purple Cluster (Statistical Methods): This smaller cluster focuses on methods highlights a technical side of modern financial analysis, stressing the value in understanding dependencies across asset classes.

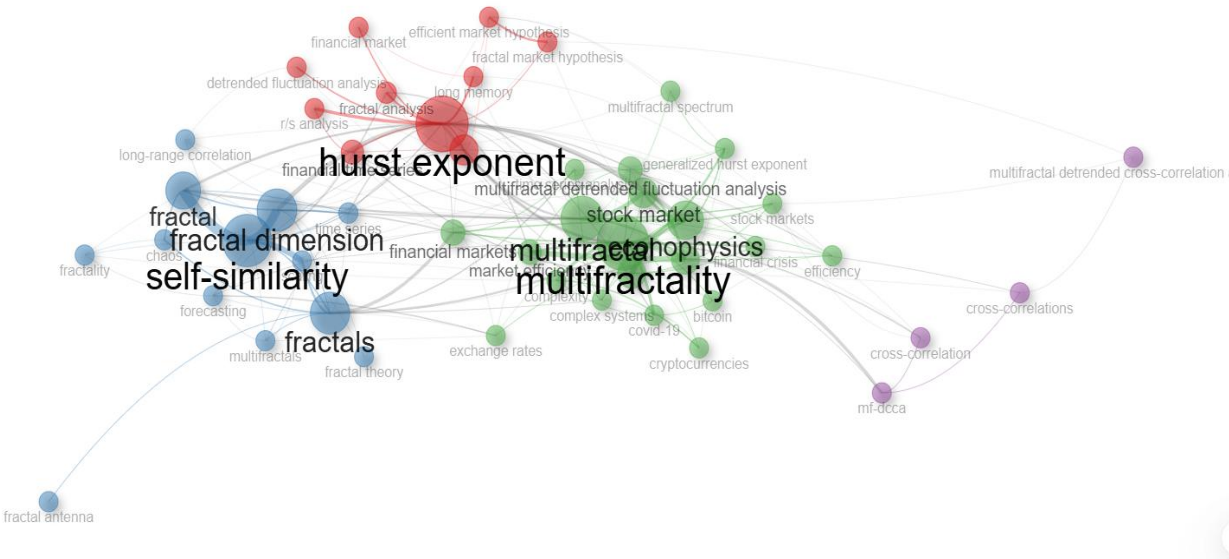

Figure 7. Co-occurrence network

The thematic map (Figure 8) further stratifies these topics by centrality and density. “Multifractality” and “multifractal analysis” emerge as motor themes that are located in the top-right quadrant with high centrality/ density. They remain the methodological and theoretical backbone of the discipline, fundamental to modelling (and thus having potential predictive power over) long-memory and volatility clustering, major cornerstones of FMH. The large number of articles identified as belonging to the multifractal domain might very well have, in themselves, provided strong bibliometric evidence for its role as a main analytical engine of FMH. M(ulti)F(range), representing at least seven measures–the possibility of picking and choosing conveniently from these descriptions, also helps make up for the lack of theoretical clarity underpinning this subdomain. As the key tool for quantifying the long-memory, heterogeneity, and scaling behavior, the FMH posits as fundamental to market dynamics, offering a more nuanced explanation than the EMH ever could.

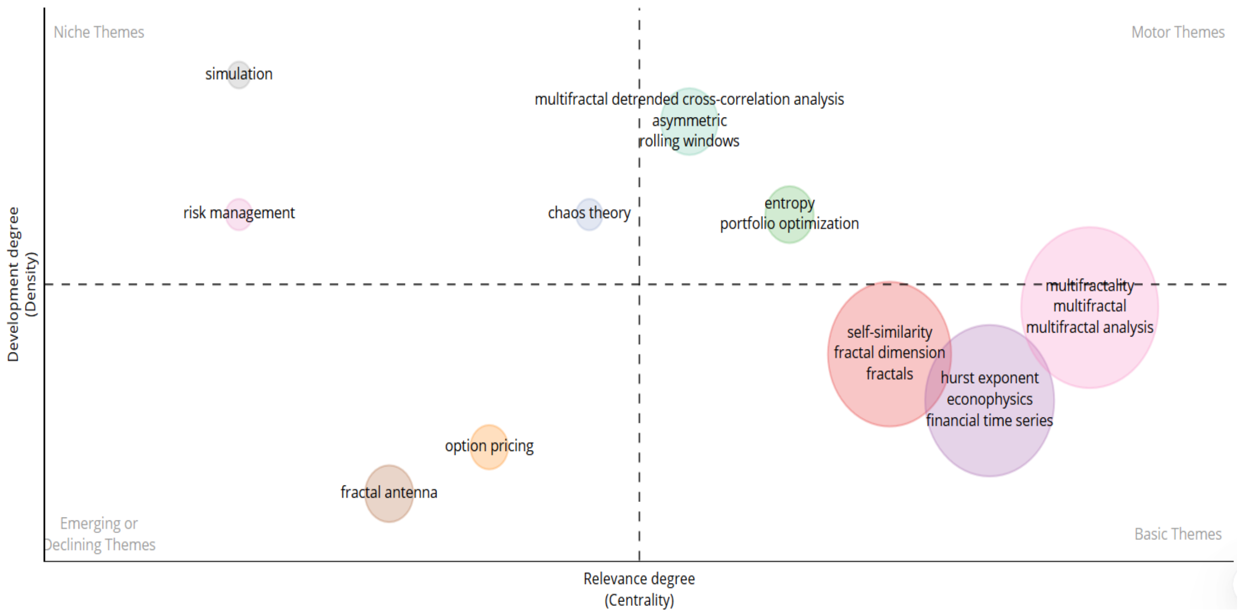

Figure 8. Thematic map

Lastly, the factorial analysis (Figure 9) supports this structure also, indicating a core cluster of applied methods (MF-DFA, cryptocurrencies), a theoretical cluster (EMH, complex systems), and an applied satellite branch (forecasting). With MF-DFA being the dominant method in most of these analyses, it is consistently applied throughout this analysis.

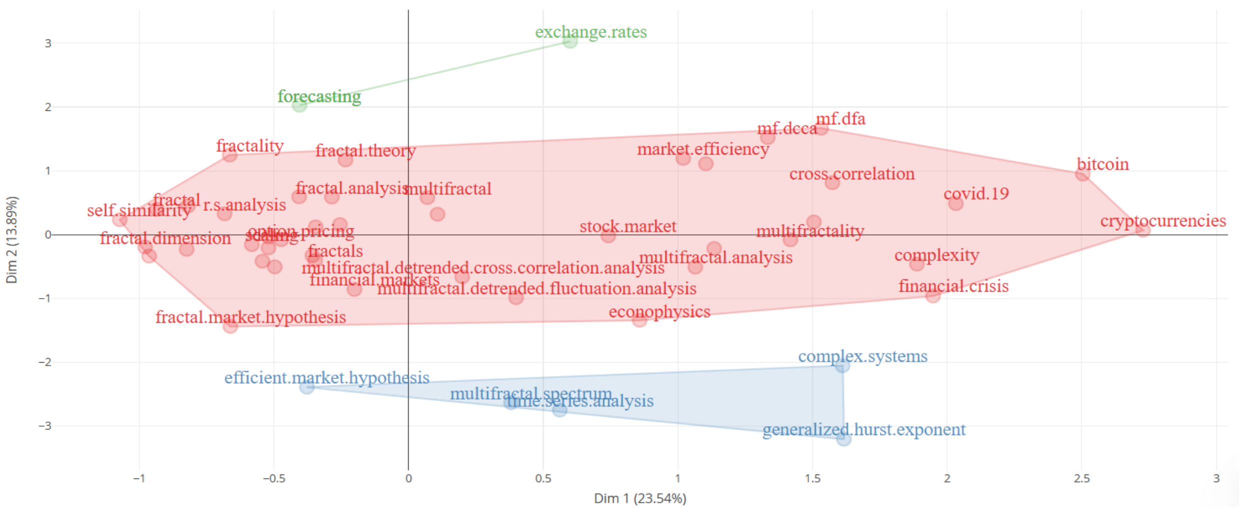

Figure 9. Factorial analysis

RQ4: Based on bibliometric evidence, what knowledge gaps exist, and what future directions are suggested to advance interdisciplinary and applied research in FMH?

Although the bibliometric evidence identified growth in papers that validate FMH, it also showed persistent gaps in our knowledge that must be addressed to move the theory beyond descriptive towards predictive. In the thematic evolution (Figure 7) and cluster analysis (Figure 3), in which we chart changes over time, surveyed studies of crises and cryptocurrencies dominated, but comparative studies between emerging and developed markets remained under-researched. Although research is focused on specific countries, the existing studies fail to generalize the applicability of FMH across economies with varying structures.

Furthermore, the shift toward analyzing systemic shocks (e.g., COVID-19) underscores a second gap; the dynamic impact of these shocks on fractal structures is not fully understood. Most studies are retrospective, limiting their predictive utility. There is a need for real-time, predictive models that can anticipate how fractal patterns will evolve during a crisis.

Lastly, the rise of terms like "deep learning" and "DeFi" suggests yet another gap; there is still a gaping void in practical use, building out trading strategies based on these fractal implications. Although many theoretical models exist, they have not been integrated into machine learning for real-time forecasting and risk management on the ground. Hence, future research should seek to: (1) conduct comparative cross-market analysis; (2) predictive modeling of fractal dynamics in systemic shocks; and (3) hybridization models for actionable financial strategies able to combine fractal indicators with artificial intelligence. Eliminating these gaps will be instrumental in pushing the FMH from a mere descriptive structure to an actual predictive and prescriptive mechanism of current finance.

5. Conclusion and suggestions

This bibliometric analysis provides strong evidence that fractal analysis has been established as an influential paradigm of financial research. Especially noteworthy in this respect has been the development and popularity of multifractal detrended fluctuation analysis (MF-DFA). The identification of multifractality as a motor theme emphasizes not only its methodological importance but also its theoretical role in the approach to long memory and volatility clustering characterizing financial time series (Jiang et al., 2019; Zhou et al., 2024).

The results of our work are in agreement with the previous reviews in economics and complex finance that have acknowledged this failure to describe financial markets from a Gaussian perspective (Peters, 1989; Evertsz, 1995). Recent works indeed have indicated that systemic shocks disrupt the fractal dynamics significantly, which is also in line with our finding on the prominent timescale of inflation, crises and cryptocurrencies in recent literature (Okorie and Lin, 2021; Zhou et al., 2022). On a theoretical basis, the strong multifractality leads to providing empirical evidence against the efficient market hypothesis and strongly supports the fractal market hypothesis (FMH) as a more valid proposition. This validates that market fractality is dynamic, not static, across different investment horizons, and that the observation of investment horizon heterogeneity is the primary motivation of FMH. The amalgamation of fractal methodologies with econophysics, complexity approaches and AI reveals the versatility of the FMH beyond conventional finance.

Fractal research provides useful practical tools from predictive disciplines and applications such as risk management, algorithmic trading, etc. The theoretical rigor of fractal-based systems alongside noise from data and cost-related challenges has limited their real-time applicability in near-scale machine-learning scenarios. Despite the challenges, there is a way forward.

6. Research gaps and future directions

Bibliometric evidence here demonstrates the previous expansion of knowledge, yet also highlights gaps in current understanding. There are several key immediate areas that future work should build on, including:

Systematic comparison of across and between developed, emerging, and frontier markets to test the universality of FMH. These studies are important to detect structural differences in fractal behavior and promote the generalizability of the theory.

Combine fractal indicators with AI using metrics like the Hurst exponent and machine learning, deep learning, and neural networks, allowing for greater predictive power and real-time analysis of dynamic and convoluted movements in the market, which are effectively great in volatile markets such as cryptocurrencies (Kehinde et al., 2023).

Utilizing longitudinal studies of how pandemics, policy changes, and geopolitical events affect fractal market structures to develop predictive models for systemic shocks. Such insights can guide policymakers towards designing more robust market mechanisms.

Real-time applications based on a computationally efficient fractal model with mere insight for algorithmic trading and portfolio optimization. This calls for a reduction in the computational complexity of instruments such as MF-DFA and wavelet analysis to render them practical in high-frequency trading settings (Popovska and Georgieva-Tsaneva, 2024).

The use of fractal finance should also be better defined in different areas, such as cryptocurrencies, DeFi (decentralized finance) and climate-related financial risks. This will elucidate incremental nonlinearities in different levels of market interactions obtained by combining fractal ideas with data science and behavioral economics.

For the first time, this research presents a complete bibliometric analysis of market fractal pattern literature from 1982 to 2025. Identifying trends regarding impactful contributors and the thematic evolution of the field through an analysis of 1,280 Scopus-indexed documents. The results illustrated how remarkably the discipline has diversified when compared to its previous orientation, earlier modelling work and theoretical considerations with current transformational transformations of artificial intelligence/and crisis-influenced subtopologies. The fact that this field of research continues to grow and diversify demonstrates how the fractal market hypothesis remains relevant in its ability to address these complexities within modern financial systems.

References

1. Amini, B. A., Bozorgmehrian, S. and Banitalebi, D. B. (2024). Designing a model to investigate the process of forming cluster fluctuations according to the fractal structure in financial markets, *Advances in Mathematical Finance and Applications*, 9(3), pp. 840-855 (In Persian).
2. Aria, M. and Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), pp. 959-975. https://doi.org/10.1016/j.joi.2017.08.007
3. Bayraktar, E., Poor, H. V. and Sircar, K. R. (2004). Estimating the fractal dimension of the SandP 500 index using wavelet analysis. *International Journal of Theoretical and Applied Finance*, 7(05), pp. 615-643. https://doi.org/10.1142/S021902490400258X

4. Berghorn, W., Schulz, M. T. and Otto, S. (2019). Fractal markets, frontiers, and factors. Frontiers, and Factors, University of Applied Sciences Aschaffenburg, Aschaffenburg, Germany. Available at SSRN. https://dx.doi.org/10.2139/ssrn.3463151
5. Bianchi, S. and Frezza, M. (2017). Fractal stock markets: International evidence of dynamical (in) efficiency. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(7), A. 071102. https://doi.org/10.1063/1.4987150
6. Blackledge, J. and Lamphiere, M. (2021). A review of the fractal market hypothesis for trading and market price prediction. *Mathematics*, 10(1), p. 117. https://doi.org/10.3390/math10010117
7. Brătian, V., Acu, A. M., Oprean-Stan, C., Dinga, E. and Ionescu, G. M. (2021). Efficient or fractal market hypothesis? A stock indexes modelling using geometric Brownian motion and geometric fractional Brownian motion. *Mathematics*, 9(22), p. 2983. https://doi.org/10.3390/math9222983
8. Castilhos, R. B., Dolbec, P. Y. and Veresiu, E. (2017). Introducing a spatial perspective to analyze market dynamics. *Marketing Theory,* 17(1), pp. 9-29. https://doi.org/10.1177/1470593116657915
9. Donthu, N., Kumar, S., Pandey, N. and Soni, G. (2021). A retrospective overview of Asia Pacific Journal of Marketing and Logistics using a bibliometric analysis. *Asia Pacific Journal of Marketing and Logistics*, 33(3), pp. 783-806. https://doi.org/10.1108/APJML-04-2020-0216
10. El-Nabulsi, R. A. and Anukool, W. (2025). Qualitative financial modelling in fractal dimensions. *Financial Innovation*, 11(1), p. 42. https://doi.org/10.1186/s40854-024-00723-2
11. Evertsz, C. J. (1995). Fractal geometry of financial time series. *Fractals*, 3(03), pp. 609-616. https://doi.org/10.1142/S0218348X95000539
12. Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *The Journal of Finance*, 25(2), pp. 383-417. https://doi.org/10.2307/2325486
13. Farinelli, S. (2015). Geometric arbitrage theory and market dynamics. Bonstetten, Switzerland. Available at SSRN 1113292. https://dx.doi.org/10.2139/ssrn.1113292
14. Gayathri, M., Selvam, M., Lingaraja, K., Vasanth, V. and Karpagam, V. (2013). Fractal dimension of SandP Cnx nifty stock returns. *Asian Journal of Empirical Research*, 3(9), pp. 1166-1190.
15. Goyal, R. (2024). Exploring the fractal geometry of financial time series. *Mod. Dyn. Math. Prog*, 1(1), pp. 1-5.
16. Grönlund, A., Yi, I. G. and Kim, B. J. (2012). Fractal profit landscape of the stock market. *PloS One,* 7(4), A. e33960.
17. Gu, M. (2025). Fractal structures in financial markets: evidence from cross-asset and multi-timeframe pattern similarity, International Academy Okma, Bloomfield Hills, Michigan. https://dx.doi.org/10.2139/ssrn.5251815
18. Huang, B. N. and Yang, C. W. (1995). The fractal structure in multinational stock returns. *Applied Economics Letters,* 2(3), pp. 67-71. https://doi.org/10.1080/135048595357591
19. Jiang, Z. Q., Xie, W. J., Zhou, W. X. and Sornette, D. (2019). Multifractal analysis of financial markets: A review. *Reports on Progress in Physics*, 82(12), A. 125901. https://doi.org/10.1088/1361-6633/ab42fb
20. Karp, A. and Van Vuuren, G. (2019). Investment implications of the fractal market hypothesis. *Annals of Financial Economics*, 14(01), A. 1950001. https://doi.org/10.1142/S2010495219500015
21. Kehinde, T. O., Chan, F. T. and Chung, S. H. (2023). Scientometric review and analysis of recent approaches to stock market forecasting: Two decades survey. *Expert Systems with Applications*, 213(1), A. 119299. https://doi.org/10.1016/j.eswa.2022.119299
22. Ko, P. S. and Kuo-Shing, C. (2025). Discovering AI tokens in the fractal markets hypothesis and

their time-frequency co-movements with the leading high-carbon cryptocurrency. *Data Science in Finance and Economics,* 5(3), p. 293. https://doi.org/10.3934/DSFE.2025013
23. Kristjanpoller, W. and Tabak, B. M. (2025). Day of the week effect on the crypto market: A high-frequency asymmetric multifractal analysis. *Physica A: Statistical Mechanics and its Applications*, 658, A. 130306. https://doi.org/10.1016/j.physa.2024.130306
24. Leventides, J., Melas, E., Poulios, C., Livada, M., Poulios, N. C. and Boufounou, P. (2024). Application of the fractal Brownian motion to the Athens stock exchange. *Fractal and Fractional*, 8(8), p. 454. https://doi.org/10.3390/fractalfract8080454
25. Liu, S., Tu, Q., Gong, Q. and Wang, Y. (2025). A fractal and game-theoretic approach to stock market analysis: an energy-space-time framework for trading strategy optimization. *Frontiers in Economics and Management,* 6(8), pp. 41-50. https://doi.org/10.6981/FEM.202508_6(8).0005
26. Malik, P. (2011). Redesigning the stock market: A fractal approach. Sage Publications Pvt, Delhi, India.
27. Mandelbrot, B. (1972). Statistical methodology for nonperiodic cycles: From the covariance to R/S analysis. *Annals of Economic and Social Measurement*, 1(3), A. 259290.
28. Moradi, M., Jabbari Nooghabi, M. and Rounaghi, M. M. (2021). Investigation of fractal market hypothesis and forecasting time series stock returns for Tehran Stock Exchange and London Stock Exchange. *International Journal of Finance and Economics,* 26(1), pp. 662-678. https://doi.org/10.1002/ijfe.1809
29. Nourahmadi, M. (2024). Financial insights: harnessing recommender systems through bibliometric analysis. *Knowledge Economy Studies*, 1(1), pp. 93-116 (In Persian). https://doi.org/10.22034/kes.2024.2037891.1003
30. Nourahmadi, M. and Parsi, F. (2025). The role of artificial intelligence in enhancing green accounting and sustainable development: a bibliometrix method. *Empirical Research in Accounting,* 15(2), pp. 211-238 (In Persian). https://doi.org/10.22051/jera.2025.50235.3512
31. Nourahmdi, M. and Rasti, F. (2025). Shaping fintech through regulations: Insights and future directions. *Knowledge Economy Studies*, 2(1), pp. 35-57 (In Persian). https://doi.org/10.22034/kes.2025.2056916.1052
32. Okorie, D. I. and Lin, B. (2021). Stock markets and the COVID-19 fractal contagion effects. *Finance Research Letters*, 38, A. 101640. https://doi.org/10.1016/j.frl.2020.101640
33. Onali, E. and Goddard, J. (2011). Are European equity markets efficient? New evidence from fractal analysis. *International Review of Financial Analysis*, 20(2), pp. 59-67. https://doi.org/10.1016/j.irfa.2011.02.004
34. Panas, E. (2001). Estimating fractal dimension using stable distributions and exploring long memory through ARFIMA models in Athens Stock Exchange. *Applied Financial Economics*, 11(4), pp. 395-402. https://doi.org/10.1080/096031001300313956
35. Peters, E. E. (1989). Fractal structure in the capital markets. *Financial Analysts Journal*, 45(4), pp. 32-37. https://doi.org/10.2469/faj.v45.n4.32
36. Popovska, E. and Georgieva-Tsaneva, G. (2024). Fractal-based robotic trading strategies using detrended fluctuation analysis and fractional derivatives: A case study in the energy market. *Fractal and Fractional,* 9(1), p. 5. https://doi.org/10.3390/fractalfract9010005
37. Samadder, S., Ghosh, K. and Basu, T. (2013). Fractal analysis of prime Indian stock market indices. *Fractals,* 21(01), A. 1350003. https://doi.org/10.1142/S0218348X13500035
38. Verma, S. K. and Kumar, S. (2023). Fractal dimension analysis of stock prices of selected resulting companies after mergers and acquisitions. *The European Physical Journal Special Topics*, 232(7), pp. 1093-1103. https://doi.org/10.1140/epjs/s11734-023-00863-z

39. Xu, C., Ke, J., Peng, Z., Fang, W. and Duan, Y. (2022). Asymmetric fractal characteristics and market efficiency analysis of style stock indices. *Entropy*, 24(7), p. 969. https://doi.org/10.3390/e24070969
40. Yang, K. I. (2025). The 1987 crash as the fractal DNA of modern stock market crashes: a technical analysis. (pp. 1-23). Available at SSRN 5173548. https://dx.doi.org/10.2139/ssrn.5173548
41. Yu, J., Baojiang, Z. and Pengjian, S. (2006). Multifractal analysis of stock market time series. *Journal of Beijing Jiao Tong University,* 30(6), pp. 69-72.
42. Zhou, W., Huang, J. and Wang, M. (2024). Multifractal characteristics and information flow analysis of stock markets based on multifractal detrended cross-correlation analysis and transfer entropy. *Fractal and Fractional*, 9(1), p. 14. https://doi.org/10.3390/fractalfract9010014

IJAAF; Vol. 10, No. 2, Spring 2026, Serial 35

Ferdowsi University of Mashhad

Iranian Journal of
Accounting, Auditing & Finance

Quarterly

RESEARCH ARTICLE

The Impact of the COVID-19 Crisis on Corporate Performance: The Moderating Role of Cash and Human Resources

Najmeh Hajian*, **Mahdieh Yazdanbakhsh**

Department of Accounting, Khatam University, Tehran, Iran

How to cite this article:
Hajian, N. and Yazdanbakhsh, M. (2026). The Impact of the COVID-19 Crisis on Corporate Performance: The Moderating Role of Cash Resources and Human Resources. Iranian Journal of Accounting, Auditing and Finance, 10(2), 21-36. doi: 10.22067/ijaaf.2026.47547.1584
https://ijaaf.um.ac.ir/article_47547.html

ARTICLE INFO

Article History
Received: 2025-09-03
Accepted: 2025-12-10
Published online: 2026-04-12

Abstract

The main objective of this study was to examine the impact of the COVID-19 crisis on firms' financial and market performance. The study employed quarterly data from a sample of firms listed on the Tehran Stock Exchange, covering 22 quarters 11 quarters before the COVID-19 crisis and 11 quarters during the pandemic. In total, 2,310 firm-quarter observations were analyzed using panel data methods in EViews 13 software. The findings indicated that during the quarters affected by COVID-19, the pandemic had a significant negative effect on the operating profit ratio and return on assets. In contrast, due to the inflationary conditions induced by the pandemic, COVID-19 exerted a significant positive effect on stock prices and, consequently, on firms' market performance. Moreover, the results showed that adequate cash holdings mitigate the adverse effects of COVID-19 on both financial and market performance, suggesting that sufficient liquidity can serve as a financial buffer during crises. Conversely, a larger workforce amplifies the negative effect of COVID-19 on financial performance, given that COVID-19 is a human-centered crisis, while it does not influence the relationship between COVID-19 and market performance. Beyond confirming the expected negative impact of the COVID-19 crisis on firms' financial performance, the findings highlighted two novel contributions. First, they showed that maintaining adequate cash holdings can serve as a defensive shield for firms during crises. Second, they revealed that a larger workforce may intensify the adverse effects of human-centered crises, such as the COVID-19 pandemic. Accordingly, the study suggested that firms should prioritize holding sufficient cash reserves as well as implementing human resource policies, such as training, preparedness, and succession planning, to better withstand future crises.

Keywords:
Cash Resources, Corporate Financial Performance, Corporate Market Performance, COVID-19 Crisis, Human Resources

https://doi.org/10.22067/ijaaf.2026.47547.1584

CC BY

NUMBER OF REFERENCES
27

NUMBER OF FIGURES
0

NUMBER OF TABLES
7

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

*Corresponding Author: Najmeh Hajian
Email: hajiyan.najmeh@gmail.com
Tel: 09122589633
ORCID: 0009-0006-7426-4166

1. Introduction

Economic crises typically lead to a reduction in economic activities and exert significant effects on economic growth, inflation, employment, industrial production, and wholesale and retail prices. These crises, whether short-term or long-term, ultimately reduce the overall economic welfare of society (Pishbahar and Rasouli, 2019).

Crises have always impacted conventional economic frameworks and procedures. Examples of such crises include wars, floods, earthquakes, terrorist attacks, and even outbreaks of contagious diseases. One of the recent significant and influential crises affecting the global economy, including Iran's, has been the outbreak of the coronavirus (COVID-19) pandemic. The onset of COVID-19 in China and its widespread propagation in other countries exemplify a true crisis of unprecedented scale in history (Ashrafi and Kazempour, 2019). While the COVID-19 outbreak has threatened public health, its economic consequences are also highly substantial (Qin et al., 2021). Recent studies have demonstrated that exogenous shocks caused by the COVID-19 crisis have had significant negative effects on firm performance (Almustafa et al., 2023; Shen et al., 2021), corporate investment (Hajian and Hashemzadeh, 2024; Zheng and Zhang, 2021), and the efficiency of financial institutions and investment funds (Zheng and Zhang, 2021). However, to date, no study in Iran has examined the impact of this crisis on firms' financial and market performance. The findings of this research can not only enrich the existing literature but also assist companies in identifying the effects of crises and preparing for their adverse consequences. Accordingly, the first research question is formulated as follows:

Does the COVID-19 crisis affect the financial and market performance of firms listed on the Tehran Stock Exchange?

On the other hand, one of the factors that can act as a defensive shield to protect firms from the financial impacts of various crises is the availability of sufficient cash resources or financial flexibility. Cash flexibility refers to a company's ability to secure financial resources to appropriately respond to unforeseen events in order to maximize firm value (Byoun, 2011). Modern business models require sufficient flexibility to adapt to crisis conditions. Companies that are able to continue their economic activities are those that can restructure and adapt their operations to the surrounding environmental conditions (Schweizer and Nienhaus, 2017). The importance of adequate cash reserves becomes particularly evident during financial crises such as the COVID-19 pandemic, as they help reduce financial constraints caused by lower investment, increased financing burdens, and the potential loss of future investment opportunities. During high financing costs, holding cash can be a buffer for companies to strategically secure their profit margins against the negative impact of financial crises (Zheng, 2022). Nevertheless, excessive cash reserves may also increase the agency costs of shareholder interests. These costs are a consequence of holding too much liquidity, which reflects the downside of maintaining excess liquidity. Therefore, firms should strike an appropriate balance between holding and utilizing cash during crises. Cash holdings, therefore, represent a double-edged issue. On the one hand, they may mitigate the adverse effects of a crisis, while on the other hand, they may signal inefficient use of financial resources and lead to weaker performance. Examining this issue can thus provide valuable insights for firms. Accordingly, the second research question is formulated as follows:

Do higher levels of cash holdings mitigate the negative effects of the COVID-19 crisis on firm performance?

Another critical factor that may influence the extent to which firms are affected by crises, particularly human-driven crises such as pandemics, is human resources. The level of dependence on human capital varies by company based on factors such as business type, technological intensity, operational modernization, labor- vs. capital-intensiveness, and other determinants. The greater the number of employees in a company, the more pronounced human-related challenges and crises tend to become. Although human resource challenges are diverse, they became more significant with the onset of the COVID-19 crisis (Onwuegbuna et al., 2021). Researchers such as Carnevale and Hatak (2020) have examined the prominent role of human resources in shaping how firm performance is affected by COVID-19. Jetter et al. (2020) and Mankin et al. (2020) argue that economic crises and unforeseen events generally reduce profitability and financial performance across various types of businesses, with the most severe impacts borne by companies with large workforces. Therefore, in contrast to the expectation that higher levels of cash reserves would mitigate the negative impacts of the COVID-19 crisis, a greater reliance on human resources appears to intensify the adverse effects of the crisis.

Therefore, greater reliance on human resources may intensify the adverse effects of crises with human origins. Iran, as one of the countries that experienced relatively high rates of mortality and illness during the COVID-19 pandemic, provides a relevant context to examine how firms' human resources affected their performance during this crisis. This issue not only highlights the importance of the human resource dimension but can also guide firms in managing future crises. Accordingly, the third research question is formulated as follows:

Do higher levels of human resources amplify the negative effects of the COVID-19 crisis on firm performance?

The rest of this paper proceeds as follows. The next section reviews the theoretical background and related literature, which leads to the formulation of six research hypotheses. Section three describes the research methodology, including the regression models and variables. Section four presents the descriptive statistics and reports the empirical findings. Section five provides a discussion of the results and concluding remarks. The final section outlines the study's implications and acknowledges its limitations.

2. Literature review

2.1. COVID-19 Crisis and Corporate Performance

In recent years, the world has witnessed the COVID-19 pandemic as one of the most severe crises in human history, which led to widespread economic devastation and caused global unrest. The disease, because of extreme contagion and virulence, forced lockdowns at global ports and airports and strict labor protocols for fear of spreading the ailment from person to person or through the air, as well as restrictions on imports and exports. Preventive measures taken around the world caused many businesses to close, creating supply and demand imbalances that contributed to serious economic difficulties and recession (Ding et al., 2021).

As news of the outbreak first began to emerge, fears of its deep economic impact gripped the world. Social distancing, self-isolation and travel restrictions reduced the availability of labor across all sectors of the economy and caused massive job losses. Schools and universities were shuttered, while demand for many goods and manufactured products fell off. On the other hand, there was a strong surge in medical supply and food products demand, which was mainly due to stockpiling behavior (Dhar et al., 2020). At the macro level, COVID-19 sparked the deepest global recession since the Great Depression. Among the pandemic's most negative consequences were declines in gross domestic product (GDP), widespread business bankruptcies, and massive job losses. However, how did COVID-19 impact corporate performance?

A 2021 Deloitte study of the COVID-19 crisis found a significant effect on firms' internal operations. This research is an obvious example of the detrimental impact that the pandemic has on internal performance in companies. According to Deloitte's global survey conducted in the first quarter of 2020, a staggering 80% of corporate executives noted that COVID-19 negatively affected their companies' performance. The key reasons reported for this adverse effect included falling demand, operating restrictions, an inability to plan because of increased uncertainty, and changes in consumer behavior. This study demonstrates how an extraordinary and unforeseen event can cause an extreme disruption to the internal corporate operation.

In fact, the pandemic was influencing performance through multiple channels, both direct and indirect. In the short term, companies experienced waning revenues owing to a decline in consumer demand, interruptions to the supply chain, operational challenges arising from shutdowns and labor shortages and increased expenses related to public health and safety protocols. Indirect pressures from greater uncertainty and volatility in financial markets also had the effect of limiting firms' access to capital for investment and reducing overall confidence in markets. In addition, changes in consumer habits—from the acceleration of e-commerce to increased spending on essential goods—led to heterogeneous effects across sectors, with some companies thriving and others suffering losses.

Additionally, Acharya and Steffen (2020), researchers at Ohio State University, asserted that the COVID-19 crisis negatively impacts corporate performance. Their research explored how the pandemic helped to drive declines in profitability and financial performance across the board. The results showed that COVID-19 resulted in significant adverse impacts, such as revenue losses, higher costs, and reduced production capacities.

Additionally, Shen et al. (2021) studied the effects of the COVID-19 crisis on the performance of Chinese firms. The pandemic, their study also noted, negatively affected firm performance, especially for the companies with low investment capacities and lower-income levels.

Looking at the analysis of previous research in this area, the following research hypotheses are proposed:

H1: There is a significant relationship between the outbreak of COVID-19 and the financial performance of companies.

H2: There is a significant relationship between the outbreak of COVID-19 and the market performance of companies.

2.2. COVID-19 crisis and cash holdings

Crises represent exogenous shocks to economic activities, which implies that such a shock is beyond the control of managers. Therefore, organizations need strategies to enhance their resilience to mitigate the destructive effects of such crises. A severe human calamity like the COVID-19 pandemic creates various limitations and challenges for firms to face, such as reduced consumer demand, operational disruptions, planning uncertainty, changes in consumption patterns, supply chain collapse and increased operating costs, all of which further lead to a downward spiral of corporate performance. Adequate cash reserves or financial institutionalization ability is one of the crucial factors that can protect firms from the consequences of crises (Zheng, 2022).

Cash flexibility is defined as a firm's capacity to source the best available financial resources to react optimally to each contingency event maximizing firm value (Byoun, 2011). Modern business models require this flexibility to adapt and restructure operations in response to changing environmental conditions (Schweizer and Nienhaus, 2017). The importance of adequate cash reserves becomes especially clear during financial crises like the COVID-19 pandemic, as they help alleviate financial constraints caused by reduced investment, increased financing burdens, and potential loss of future investment opportunities. The advantage of keeping a liquid balance sheet is vital for a business to stay alive in bad times. In theory, financing frictions combined with deteriorating market conditions can lead to credit rationing and exacerbate refinancing risk, making externally obtained capital more expensive than that generated internally (Zheng, 2022).

One of the leading theories regarding corporate cash balances is the Trade-off Theory. Under this theory, firms decide on the optimal amount of cash holdings by looking at the trade-off between the benefits and costs of holding cash. There are three main reasons for holding cash, including the transactional motive, the precautionary motive and the speculative motive. According to the precautionary motive, companies accumulate cash in times of high uncertainty to limit the risk of financial distress. The precautionary motive of corporate cash holds a value to firms in the bad states when firms face credit tightening and lack access to external finance (Zheng, 2022). Given that crises tend to create short-term uncertainties in business operations, maintaining a sufficient level of liquidity allows firms to cope with emergencies without dependence on external financing (Opler et al., 1999). Additionally, since some firms are more constrained than others in their ability or the cost of external financing, managers are encouraged to build up cash reserves before a contraction in credit leads to cash flow deficits. Having a prudent cash buffer enables firms to mitigate refinancing risk. In this context, Zheng (2022) and Chiappini et al. (2021) highlight the importance of corporate cash holdings in mitigating the negative effects of the COVID-19 crisis. The COVID-19 shock thus presents a unique opportunity to investigate the influence of unforeseen crises on firm performance and to assess the importance of cash reserves in shaping this impact. This issue is addressed in two hypotheses of the study as follows:

H3: The level of corporate cash holdings moderates the relationship between the COVID-19 crisis and financial performance.

H4: The level of corporate cash holdings moderates the relationship between the COVID-19 crisis and market performance.

2.3. COVID-19 Crisis and human resources

Just as the financial crisis of 2007–2009 highlighted the role of chief financial officers, the COVID-19 pandemic brought human resource (HR) managers to the forefront (Caligiuri et al., 2022) because HR management focuses on overseeing a unique resource, human capital, which is fundamentally different from financial or physical resources. While other resources enable work to be carried out, it is human resources that actually perform the work (Opatha, 2020).

COVID-19 is considered a human-centric crisis, primarily affecting people more than any other resource, thereby disrupting both organizational operations and HR functions. Although human resource challenges are diverse, the emergence of the COVID-19 pandemic has significantly intensified them (Onwuegbuna et al., 2021). Mwita (2020) found that during the COVID-19 crisis, recruitment and selection processes drastically slowed, employee training programs were canceled either to protect employees from the virus or due to limited online infrastructure, and performance management became increasingly difficult. In some organizations, financial obligations and benefit payments were also suspended. As a result, organizations were confronted with unprecedented and unconventional situations for which many of their employees had neither training nor expectations.

Although crises, particularly those with human-centric origins, such as the COVID-19 pandemic, can significantly affect employee well-being, mental health, and productivity, thereby influencing overall organizational performance, their impact appears to be more pronounced in labor-intensive firms. In such organizations, the heavy reliance on human capital makes them more vulnerable to disruptions affecting the workforce. While capital-intensive or highly mechanized firms are not entirely immune to these effects, the integration of technology and automated systems in their operations may buffer the severity of the crisis's impact. Consequently, the adverse effects of COVID-19 on performance are expected to be less severe in firms with lower human resource dependence. For example, Carnevale and Hatak (2020) investigated the role of human resources in mediating the effects of external crises and unforeseen events on organizational performance. Furthermore, Jetter et al. (2020) and Mankin et al. (2020) found that the most significant declines in financial performance during economic crises were observed in firms with larger workforces.

Based on these insights, this study formulates the fifth and sixth hypotheses to explore the moderating role of human resources in the relationship between the COVID-19 crisis and firm performance. It is expected that firms with larger workforces may experience amplified negative effects of the COVID-19 pandemic on their performance.

H5*:* The number of human resources moderates the relationship between the COVID-19 crisis and firms' financial performance.

H6*:* The number of human resources moderates the relationship between the COVID-19 crisis and firms' market performance.

3. Methodology

This descriptive and post-event study is a retrospective research based on actual data extracted from financial statements. The statistical population includes all companies listed on the Tehran Stock Exchange. The accessible population for hypothesis testing was selected based on the following criteria:

1. Availability of complete financial statements;
2. Exclusion of companies in the financial sector, including banks, investment companies, holdings, insurance, and leasing firms, due to differences in financial information disclosure.
3. No change in fiscal year during the study period.
4. Fiscal year ending on March 20th.
5. No trading suspension of more than three months during the study period.

After applying the above restrictions, the final research sample includes 105 companies observed over 22 quarters (2,310 firm-quarters) from 2017 to 2022. Considering that the COVID-19 crisis began in late December 2019 and its serious effects continued until late September 2022, the 11 quarters within this period were considered the "COVID-19 period" and the 11 quarters prior as the "pre-COVID-19 period."

Part of the variable data was collected from secondary sources, including various financial databases, software platforms, and websites related to the Tehran Stock Exchange. The other part, concerning human resource variables, was manually extracted from the companies' board of directors' activity reports.

3.1. Empirical models

Based on the six research hypotheses and the two financial performance metrics, a total of nine regression models are examined as follows:

Regression equations for Hypotheses 1 and 2:

1. $OP = \alpha_1 + \alpha_2 COVID19 + \alpha_3 SIZE + \alpha_4 CF + \alpha_5 LEV + \alpha_6 BOI + \alpha_7 AGE + \alpha_8 INS + \alpha_9 FAR$
2. $ROA = \alpha_1 + \alpha_2 COVID19 + \alpha_3 SIZE + \alpha_4 CF + \alpha_5 LEV + \alpha_6 BOI + \alpha_7 AGE + \alpha_8 INS + \alpha_9 FAR$
3. $TQ = \alpha_1 + \alpha_2 COVID19 + \alpha_3 SIZE + \alpha_4 CF + \alpha_5 LEV + \alpha_6 BOI + \alpha_7 AGE + \alpha_8 INS + \alpha_9 FAR$

Regression equations for Hypotheses 3 and 4:

1. $OP = \alpha_{1+} \alpha_2 COVID19{*}CH + \alpha_3 COVID19 + \alpha_4 CH + \alpha_5 SIZE + \alpha_6 CF + \alpha_7 LEV + \alpha_8 BOI + \alpha_9 AGE + \alpha_{10} INS + \alpha_{11} FAR$
2. $ROA = \alpha_{1+} \alpha_2 COVID19{*}CH + \alpha_3 COVID19 + \alpha_4 CH + \alpha_5 SIZE + \alpha_6 CF + \alpha_7 LEV + \alpha_8 BOI +$

$\alpha_9AGE + \alpha_{10}INS + \alpha_{11}FAR$

3. $TQ = \alpha_1 + \alpha_2COVID19*CH + \alpha_3COVID19 + \alpha_4CH + \alpha_5SIZE + \alpha_6CF + \alpha_7LEV + \alpha_8BOI + \alpha_9AGE + \alpha_{10}INS + \alpha_{11}FAR$

Regression equations for Hypotheses 5 and 6:

1. $OP = \alpha_1 + \alpha_2COVID19*N + \alpha_3COVID19 + \alpha_4N + \alpha_5SIZE + \alpha_6CF + \alpha_7LEV + \alpha_8BOI + \alpha_9AGE + \alpha_{10}INS + \alpha_{11}FAR$
2. $ROA = \alpha_1 + \alpha_2COVID19*N + \alpha_3COVID19 + \alpha_4N + \alpha_5SIZE + \alpha_6CF + \alpha_7LEV + \alpha_8BOI + \alpha_9AGE + \alpha_{10}INS + \alpha_{11}FAR$
3. $TQ = \alpha_1 + \alpha_2COVID19*N + \alpha_3COVID19 + \alpha_4N + \alpha_5SIZE + \alpha_6CF + \alpha_7LEV + \alpha_8BOI + \alpha_9AGE + \alpha_{10}INS + \alpha_{11}FAR$

Dependent variables

A.Financial performance:

This study adopted a model proposed by Zheng (2022), using the following two criteria to assess firm performance.

Operating Profit Ratio (OP): Calculated by dividing operating profit before depreciation by total assets.

Return on Assets (ROA): Defined as the net income of the company divided by the total assets of the company.

B. Market-based performance:

Tobin's Q Ratio (TQ): Obtained by dividing the market value of assets by the book value of assets. The market value of assets is calculated as Book Value of Liabilities + Market Value of Shareholders' Equity (Zheng, 2022; Hajian and Hashemzadeh, 2024).

Independent variable

COVID-19 Crisis (COVID-19): A dummy variable that takes the value 0 for the 11 quarters prior to the COVID-19 crisis (Spring 2017 to Fall 2019), and the value 1 for the 11 quarters during the COVID-19 period (Winter 2019 to Summer 2022) (Zheng, 2022; Hajian and Hashemzadeh, 2024).

Moderating variable for hypotheses 3 and 4:

Cash Holdings (CH): In line with previous research, cash holdings are calculated as the sum of cash and cash equivalents (as reported in the balance sheet) divided by total book assets (Zheng, 2022).

Moderating variable for hypotheses 5 and 6:

Number of Employees (N): Refers to the number of human resources reported annually in the Board of Directors' activity report (Carnevale and Hatak, 2020).

Control variables

Firm Size (SIZE): The natural logarithm of the total book value of the firm's assets (Zheng, 2022).

Operating Cash Flow (CF): Operational cash flows as reported in the statement of cash flows (Zheng, 2022).

Financial Leverage (LEV): Total liabilities divided by total assets of the company (Zheng, 2022).

Board Independence (BOI): The ratio of non-executive board members to total board members (Zandi Gargari and Amirhosseini, 2016).

Firm Age (AGE): The natural logarithm of the company's age from the year of establishment to the end of the study period (Hajian and Hashemzadeh, 2024).

Ownership Concentration (INS): Sum of the ownership percentages held by shareholders owning more than 5% of the company's shares (Hajian and Hashemzadeh, 2024).

Fixed Asset Ratio (FAR): Ratio of fixed assets to total assets (Zandi Gargari and Amirhosseini, 2016).

4. Findings

4.1. Descriptive statistics

Table 1 presents the descriptive statistics for the quantitative and qualitative variables examined in the studied companies.

Table 1. Descriptive statistics

Variable Indicator	Description	Sample Size	Mean	Median	Max	Min	Std. Dev.
OP	Operating Profit Ratio (%)	2310	0.440	0.030	18.360	-68.340	4.310
ROA	Return on Assets (%)	2310	0.360	0.020	19.180	-60.630	3.800
TQ	Tobin's Q	2310	9.600	4.190	1581.440	0.550	57.460
CH	Cash Holdings	2310	0.090	0.050	0.930	0.000	0.120
N	Number of Employees	2310	901.950	343.000	171880	22.000	2104.940
SIZE	Firm Size (log of total assets)	2310	12.910	12.830	15.330	11.080	0.640
CF	Operating Cash Flow	2310	0.060	0.040	0.750	-0.500	0.100
LEV	Financial Leverage	2310	0.550	0.580	0.980	0.020	0.210
BOI	Board Independence (%)	2310	66.440	60.000	100.000	0.000	18.160
INS	Ownership Concentration (%)	2310	68.470	72.020	99.000	0.000	18.680
FAR	Fixed Asset Ratio	2310	0.210	0.150	0.920	0.000	0.180

The dataset used in this study is quarterly, which is particularly relevant for interpreting performance-related variables. For example, the average quarterly return on assets (ROA) is approximately 0.36%, whereas the maximum observed value reaches around 19%. Descriptive statistics further reveal that the number of employees among the companies varies widely, ranging from 22 to 17,188 employees. Moreover, the average Tobin's Q ratio, which reflects the market value relative to the book value of the firms, is approximately 9.6, indicating that, on average, market value is more than nine times the book value.

4.2. Research findings

All research data have been evaluated for suitability in conducting inferential statistical analyses. To test for normality, the Jarque–Bera test was applied. The results showed that the p-values for all variables exceeded the 5% threshold, indicating no significant deviation from normality for any variable.

Spearman's rank correlation method was carried out to evaluate the extent of association between variables. Generally, the results showed no strong correlations between independent variables, which suggests that multicollinearity was not a problem here.

We analyzed heteroscedasticity with White's test, ensuring that the final regression estimations were robust against this bias, which is indicated by the presence of heteroscedasticity in our error terms. To correct for this, White–Huber robust standard errors were used.

Also, the Durbin–Watson statistic was used to assess any autocorrelation in residuals. The Durbin–Watson values for all the regression models were between 1.5 and 2.5, so it was concluded that no autocorrelation existed.

Nine panel data regression models were estimated to test research hypotheses. The outcomes of these estimations, which also included three firm-specific performance indicators and two moderating variables, are shown in the following section.

H1 and 2:

Table 2 presents the results of regression models 1 to 3, corresponding to Hypotheses 1 and 2.

Table 2. Regression results for hypotheses 1 and 2

Variable Name	Coefficient (Model 1: OP)	Sig. (Model 1)	Coefficient (Model 2: ROA)	Sig. (Model 2)	Coefficient (Model 3: QT)	Sig. (Model 3)
COVID19	-0.380	0.010	-0.360	0.010	0.710	0.000
SIZE	3.270	0.000	2.800	0.000	-2.800	0.000
CF	0.480	0.690	0.280	0.790	6.850	0.000
LEV	0.920	0.230	0.870	0.200	-2.160	0.000
BOI	-0.000	0.710	-0.000	0.740	-0.010	0.230
AGE	-4.550	0.180	-3.220	0.310	-7.100	0.050
INS	-0.000	0.470	-0.000	0.480	0.000	0.830
FAR	0.150	0.850	0.310	0.680	-4.950	0.000
Intercept (C)	-36.250	0.000	-32.020	0.000	59.870	0.000
R^2	0.210		0.210		0.470	
Adj. R^2	0.170		0.170		0.440	
F-stat	4.660		4.550		14.970	
P-value	0.000		0.000		0.000	
DW Stat.	2.460		2.470		1.620	

The F-statistics for the panel data regression models are 4.66, 14.97, and 4.55, respectively, indicating the overall significance of all three regression models. As shown at the bottom of Table 2, the R-squared values for the models are 21%, 47%, and 21%, respectively. Therefore, in the specified regression equations, approximately 21%, 47%, and 21% of the variations in the operational performance measures (Operating Profit Ratio, Tobin's Q Ratio, and Return on Assets, respectively) of the companies under study are explained by the independent and control variables included in the models.

The results presented in Table 2 show that the significance levels for the relationship between the COVID-19 crisis and the three dependent variables—Operating Profit Ratio, Return on Assets, and Tobin's Q are 0.01, 0.01, and 0.00, respectively. All of these are lower than the 5% significance level used in this study. Thus, at the 95% confidence level, Hypothesis 1 (the impact of COVID-19 on accounting-based performance) and Hypothesis 2 (the effect on market-based performance) are statistically supported.

One significant finding from this part is that the COVID-19 crisis has a significant detrimental impact on both of the accounting-based performance measures, which are the Operating Profit Ratio and Return on Assets. By contrast, the crisis was associated with a significant positive impact in terms of firms' market-based performance as measured by Tobin's Q.

While firms made less money during the COVID-19 period, this means market participants may have had more optimistic or forward-looking expectations, which put upward pressure on the valuation of the market despite weak financial results.

Hypotheses 3 and 4:

Table 3 presents the results of regression models 4 to 6, corresponding to Hypotheses 3 and 4.

Table 3. Regression results for hypotheses 3 and 4

Variable Name	Coefficient (Model 1: OP)	Sig. (Model 1)	Coefficient (Model 2: ROA)	Sig. (Model 2)	Coefficient (Model 3: QT)	Sig. (Model 3)
Covid19 * CH	6.200	0.000	5.450	0.000	-3.200	0.020
COVID19	-0.930	0.000	-0.820	0.000	0.740	0.140
CH	-5.030	0.000	-4.510	0.000	0.490	0.720
SIZE	2.840	0.000	2.430	0.000	-3.110	0.000
CF	0.630	0.610	0.600	0.580	7.390	0.000
LEV	1.290	0.070	1.150	0.070	-2.780	0.000
BOI	-0.010	0.19	-0.000	0.200	-0.020	0.000
AGE	-3.590	0.520	-2.840	0.560	-8.500	0.420
INS	0.000	0.660	0.000	0.650	0.000	0.310
FAR	0.150	0.850	0.290	0.670	-5.450	0.000
C (Constant)	-31.950	0.000	-27.780	0.000	66.000	0.000
R-squared	0.190		0.190		0.490	
Adjusted R-squared	0.140		0.140		0.460	
F-statistic	3.610		3.540		14.380	
Prob(F-statistic)	0.000		0.000		0.000	
Durbin-Watson stat.	2.400		2.410		1.680	

According to Table 3 and the corresponding models, the significance levels for the relationship between the COVID-19 crisis and firm performance (Operating Profit Ratio and Return on Assets) as well as market performance (Tobin's Q), considering the moderating variable of cash holdings, are 0.00, 0.00, and 0.02, respectively. These values are below the 5% significance level adopted in this study. Therefore, at the 95% confidence level, the null hypothesis—stating that cash holdings do not moderate the effect of the COVID-19 crisis on firm performance is rejected, and the main research hypothesis is confirmed. Cash holdings weaken the negative relationship between the COVID-19 crisis and financial performance, and weaken the positive relationship between the COVID-19 crisis and the market-based performance of the firms.

Hypotheses 5 and 6:

Table 4 presents the results of regression models 7 to 9, corresponding to Hypotheses 5 and 6.

According to Table 4 and the aforementioned models, the significance levels between the COVID-19 crisis and the company's financial performance (Operating Profit Ratio and return on assets), as well as the company's market performance (Tobin's Q ratio), considering the interaction coefficient of the moderating variable number of human resources—are -0.01, -0.03, and -0.95, respectively. These values indicate statistical significance for the two financial performance variables, except for Tobin's Q. Therefore, at a 95% confidence level, Hypothesis 5, which posits that the number of human resources moderates the impact of COVID-19 on firm performance, is confirmed. Specifically, given the negative coefficients, a greater number of human resources intensifies the negative relationship between the COVID-19 crisis and the company's financial performance. Thus, the main hypothesis for Tobin's Q is rejected, indicating that the number of human resources does not significantly moderate the relationship between the COVID-19 crisis and market performance.

Table 4. Regression results for hypotheses 5 and 6

Variable Name	Coefficient (Model 1: OP)	Sig. (Model 1)	Coefficient (Model 2: ROA)	Sig. (Model 2)	Coefficient (Model 3: QT)	Sig. (Model 3)
Covid19 * Nit	-0.010	0.010	-0.030	0.000	-0.950	0.310
COVID19	-0.190	0.030	0.520	0.000	-0.190	0.030
N	-0.000	0.000	-9.120	0.600	-0.000	0.000
SIZE	1.380	0.000	-2.780	0.000	1.210	0.000
CF	1.140	0.520	8.620	0.000	0.920	0.250
LEV	1.980	0.140	-1.780	0.000	1.700	0.000
BOI	0.000	0.460	-0.010	0.090	0.000	0.660
AGE	0.580	0.360	-3.590	0.120	0.570	0.380
INS	-0.000	0.430	-0.000	0.870	-0.000	0.510
FAR	1.260	0.190	-5.600	0.000	1.160	0.040
C	-20.180	0.000	52.790	0.000	-17.610	0.000
R-squared	5%		46%		4%	
Adjusted R-squared	4%		43%		4%	
F-statistic	12.250		16.400		12.020	
Significance (P-Value)	0.000		0.000		0.000	
Durbin-Watson Statistic	2.160		1.750		2.170	

5. Discussion and conclusion

According to economic experts, the risk posed by the coronavirus is significant enough to be considered the onset of a global economic recession. The analysis examines whether the COVID-19 crisis has a significant impact on company performance at both financial and market levels. Based on the results obtained, COVID-19 has a significant negative effect on firms' financial performance. The findings showed that the outbreak of the coronavirus disease has reduced companies' profitability and return on assets. These results are consistent with studies by Zheng (2022), Acharya and Steffen (2020) and Shen et al. (2021), which indicated a significant decline in corporate profitability during the COVID-19 crisis.

In contrast, the findings suggested that the COVID-19 pandemic had a significant positive influence on Tobin's Q ratio, reflecting enhanced market performance of firms. This outcome aligns with studies in developing economies, where the pandemic triggered inflationary trends that contributed to higher firm valuations and Tobin's Q ratios. For example, Ashrafi and Kazempour (2019), in a domestic study, identified positive effects of the pandemic on Tobin's Q based on expert survey responses.

However, this finding is different from the evidence observed in developed countries, such as Bose et al. (2022), who downloaded Tobin's Q for Australian firms in the COVID-19 crisis, suggesting market reactions may differ depending on economic conditions.

Across the dangerous secrets, the findings showed that the COVID-19 pandemic is detrimental to firms' financial performance while positively contributing to their market performance. This means that, on average, the stock market reacted positively to the outbreak and subsequent economic shutdowns, yielding substantial returns even as firms' financial performance withered.

In relation to cash holdings as a moderator, the results showed that cash holdings weaken the relationship between the COVID-19 crisis and firms' financial performance. That is, the incidence and spread of COVID-19 have brought less harm and a slight negative variation in the performance of firms with large cash reserves. Companies with more cash are less exposed to the fallout of the COVID-19 crisis because they're financially better positioned. These results were in line with the studies of Zheng (2022), Stephen (2020), Shen et al. (2021), and Sabouri (2022). Additionally, although firms' performance in the market had beneficial results, the inflationary conditions within the COVID-19 crisis and its negative effect on liquidity decreased these effects.

Furthermore, the results examining the moderating effect of the number of human resources on the relationship between the COVID-19 crisis and firm performance indicated that the COVID-19 outbreak exerted a more pronounced negative impact on the financial performance of firms with a higher number of employees. In other words, the results indicated that the number of human resources strengthens the negative relationship between the COVID-19 crisis and firm performance. However, this effect was not supported for the market performance variable (Tobin's ratio). It can be argued that since employees are directly affected by stress, illness, mortality, and other human-related consequences of crises such as COVID-19, firms with a larger workforce are more vulnerable to such disruptions, resulting in weaker overall financial performance. Nevertheless, market performance appears to remain unaffected by the stress, illness, or mortality of employees. Although workforce-related problems may disrupt internal operations, the firm's market performance or stock returns are not directly driven by such issues. In fact, capital market participants, both actual and potential investors, make buy and sell decisions based on observable and expected financial indicators, which ultimately determine the stock value of the firm. Investors are generally not informed about the internal workforce-related challenges caused by COVID-19, and even when informed, such factors do not significantly influence their trading decisions. Therefore, the results stated that the number of employees does not moderate the relationship between the COVID-19 crisis and firms' market performance.

The test results were consistent with Shinde et al. (2020), who demonstrated the negative effects of the COVID-19 crisis on employee performance. Further, the results aligned with Shinde et al. (2020), in which human resource management practices during the COVID-19 crisis and implications for employee welfare and job security perceptions were investigated. The high rate of psychological disorders and low resilience seen during the pandemic support these conclusions. This aligns with the findings of research by Nasirzadeh et al. (2020), which also pointed to the importance of psychological interventions.

6. Limitations and implications

This study faced several data-related limitations that may have implications for the findings. First, the research relied on quarterly financial statement data; in Iran, six-month and twelve-month financial statements are audited, whereas three-month and nine-month financial statements are not. Therefore, for the 3- and 9-month quarterly data, we had to rely on unaudited financial statements, which may introduce some measurement error in the financial performance indicators. Second, quarterly data on the number of human resources were not available, as Iranian Accounting Standard 22 (Interim Financial Reporting) does not require disclosure of human resources information in interim reports. Consequently, annual human resources data and board of directors' activity reports were used as proxies. Such a substitution could limit the accuracy of the analysis between workforce size and firm performance by potentially underestimating or overestimating the moderating role of HR on COVID-19 effects. Although these measures are the best data available, readers should interpret the findings cautiously because of these limitations.

These findings had significant implications for different interests of company stakeholders, regulatory authorities and developers of corporate strategy. Given the moderating effect of cash holdings in alleviating the impact of these gradual changes stemming from the COVID-19 crisis on firms' financial performance, we argued that managers should ensure optimal short-term liquidity levels to reduce liquidity risk exposure to unexpected crises. In this way, firms are poised to meet their debt obligations during crises and also minimize liquidity risk as much as possible. In addition, considering the role of cash holdings as a buffer moderating the negative effects of crises, regulators have to incentivize or mandate firms to hold sufficient liquidity to withstand sudden disruptions. On the other hand, considering the intensified adverse effects of human crises such as pandemics in firms with a larger workforce, it is recommended that firms with a significant number of employees implement human resource training programs, succession planning, and replacement management strategies. Such measures would enable other employees to temporarily assume the responsibilities of those affected during workforce disruptions.

7. Robustness tests

Following Zheng (2022), we conducted additional sensitivity analyses and robustness tests by examining the heterogeneous effects of COVID-19 across different industries and firm sizes. Tables 5 and 6 present regression results that incorporate interaction terms between the COVID-19 period and industry categories as well as firm size groups.

Table 5. Regression results by industry (COVID-19 Interaction Effects)

Industry	COV×IND (OP)	Sig.	COV×IND (ROA)	Sig.	COV×IND (QT)	Sig.	Interpretation
1. Pharmaceuticals and Chemicals	-0.010	0.404	0.079	0.032	-0.183	0.592	Pharmaceutical and chemical companies with significantly higher profitability (ROA) during the COVID-19 period.
2. Automotive and Parts	-0.017	0.217	0.049	0.266	-0.075	0.833	The automotive sector had an overall insignificant impact on its performance.
3. Metals and Mining	-0.008	0.642	0.092	0.047	0.246	0.550	Metal and mining companies achieved higher profitability (ROA) during the pandemic.
4. Cement, Tile, and Construction Materials	0.002	0.879	0.152	0.004	-0.498	0.223	The financial performance (ROA) of construction-related companies improved.
5. Food and Beverages	–0.029	0.025	0.016	0.655	-0.284	0.790	The operational performance (OP) of food companies weakened during the COVID-19 period.
6. Machinery and Equipment	–0.036	0.013	0.026	0.569	-0.495	0.239	The operational performance (OP) of this sector, industrial investment and equipment saw a significant decline.

Table 5 shows the regression results by industry, capturing the interaction effects of COVID-19. Other Industries, Holdings, and Agriculture (IND7) were excluded due to multicollinearity. The results highlighted heterogeneous industry-level responses to the COVID-19 pandemic. While the pharmaceutical, chemical, metal, and construction sectors experienced stronger profitability, the food and machinery industries were negatively impacted. However, the overall market valuation effect (Tobin’s Q) remained statistically insignificant for any of the examined industries.

Table 6. Regression results by firm size (COVID-19 Interaction Effects)

Firm Size Category	COV×SIZE (OP)	Sig.	COV×SIZE (ROA)	Sig.	COV×SIZE (QT)	Sig.	Interpretation
Medium Firms (SIZE2)	0.009	0.372	0.301	0.269	6.749	0.015	A positive and significant market reaction.
Large Firms (SIZE3)	–0.015	0.398	0.984	0.0505	13.662	0.000	A significant market gain during the pandemic.

Table 6 presents the regression results by firm size, capturing COVID-19 interaction effects. Firms were classified into three categories: small, medium, and large, based on their total asset size. Small firms (SIZE1) were excluded due to multicollinearity. Medium firms (SIZE2) display positive and statistically significant effects only in the market-based model (Tobin's Q), indicating that they reap benefits from optimism by investors. The results showed that large firms (SIZE3) that gained on the market in fact present smaller or insignificant OP and ROA effects. In general, the COVID-19 effects depend on firm size, while investors preferred medium and large firms in market reaction terms, market analysts relied on operational restrictions.

Pre- vs. Post-COVID analysis

A comparison was performed between the pre- and post-COVID-19 periods. In more detail, we estimated the OP, ROA, and Tobin's Q models for the 11 quarters before and after the onset of COVID-19 (i.e., onset = 01Q21). It also allows us to explore how the drivers of firm performance and market value may have shifted in light of the pandemic. The results of this split-period regression analysis are shown in Table 7. Comparing the two mentioned periods reveals significant changes in the strength and significance of several explanatory variables.

Table 7. Performance models (Post- vs. Pre-COVID-19)

Variable	OP Post-COV	Sig	OP Pre-COV	Sig	ROA Post-COV	Sig	ROA Pre-COV	Sig	QT Post-COV	Sig	QT Pre-COV	Sig
SIZE	0.020	0.008	0.035	0.000	2.829	0.000	0.548	0.000	-2.791	0.000	-2.061	0.000
CF	0.693	0.000	0.587	0.000	3.194	0.000	0.488	0.000	4.616	0.000	10.113	0.000
LEVERAGE	-0.052	0.011	-0.081	0.000	1.068	0.260	-0.070	0.290	-1.444	0.010	-0.284	0.390
BOI	0.000	0.310	0.000	0.200	-0.009	0.640	0.000	0.520	-0.030	0.060	0.033	0.008
AGE	-0.578	0.000	0.254	0.050	-8.398	0.320	-1.267	0.001	-22.714	0.000	8.260	0.030
INS	0.000	0.480	0.000	0.410	0.000	0.980	-0.001	0.340	0.032	0.040	-0.031	0.100
FAR	-0.106	0.000	-0.086	0.000	0.384	0.610	0.093	0.270	-4.211	0.000	-4.291	0.000
Intercept (C)	0.670	0.006	-0.691	0.000	-25.552	0.050	-5.640	0.000	79.097	0.000	25.516	0.000
R^2	0.567	-	0.475	-	0.211	-	0.194	-	0.494	-	0.553	-
Adj. R^2	0.521	-	0.420	-	0.127	-	0.109	-	0.440	-	0.505	-
F-stat	12.310	-	8.510	-	2.510	-	2.270	-	9.170	-	11.600	-
DW Stat	2.330	-	2.430	-	2.500	-	2.460	-	1.590	-	2.010	-

Table 7 shows that firm size and cash flow consistently remain the most significant determinants of firm performance, both before and during the COVID-19 pandemic. Larger firms and those with stronger cash flows tend to perform better and maintain higher performance levels, underscoring their resilience amid market fluctuations. Interestingly, the impact of leverage and firm age on performance has become more pronounced in the post-COVID period. This suggests that firms with higher debt levels faced increased financial strain during the crisis, while older, more mature firms demonstrated different sensitivities, potentially due to their established structures and experience in navigating turbulent environments. In contrast, other factors such as board independence, fixed asset ratio and ownership concentration exhibited limited or inconsistent effects on firm outcomes across the two periods. These results highlighted how the pandemic has intensified the importance of financial health and firm maturity while diminishing the relative influence of ownership structure on firm performance during this challenging time.

References

1. Acharya, V. V. and Steffen, S. (2020). The risk of being a fallen angel and the corporate dash for cash in the midst of COVID. *The Review of Corporate Finance Studies,* 9(3), pp. 430-471. https://doi.org/10.1093/rcfs/cfaa013
2. Almustafa, H., Nguyen, Q. K., Liu, J. and Dang, V. C. (2023). The impact of COVID-19 on firm risk and performance in MENA countries: Does national governance quality matter? *PloS One*, 18(2), A. e0281148. https://doi.org/10.1371/journal.pone.0281148
3. Ashrafi, H. and Kazempour, Z. (2019). Typology of information in the Coronavirus crisis: Presenting a perspective. *Iranian Journal of Emergency Medicine,* 6(1), pp. e24-e24 (In Persian).
4. Bose, S., Shams, S., Ali, M. J. and Mihret, D. (2022). COVID-19 impact, sustainability performance and firm value: international evidence. *Accounting & Finance*, 62(1), pp. 597-643. https://doi.org/10.1111/acfi.12801
5. Byoun, S. (2011). Financial flexibility and capital structure decision. Baylor University, Texas, Available at SSRN 1108850. https://dx.doi.org/10.2139/ssrn.1108850
6. Caligiuri, P., De Cieri, H., Minbaeva, D., Verbeke, A. and Zimmermann, A. (2022). International HRM insights for navigating the COVID-19 pandemic: Implications for future research and practice. In crises and disruptions in international business: How multinational enterprises respond to crises (pp. 417-454). Cham: Springer International Publishing, Cham, Switzerland.
7. Carnevale, J. B. and Hatak, I. (2020). Employee adjustment and well-being in the era of COVID-19: Implications for human resource management. *Journal of Business Research*, 116, pp. 183-187. https://doi.org/10.1016/j.jbusres.2020.05.037
8. Chiappini, H., Vento, G. and De Palma, L. (2021). The impact of COVID-19 lockdowns on sustainable indices. *Sustainability*, 13(4), A. 1846. https://doi.org/10.3390/su13041846
9. Dhar, B. K., Ayittey, F. K. and Sarkar, S. M. (2020). Impact of COVID-19 on Psychology among University Students. *Global Challenges,* 4(11), A. 2000038. https://doi.org/10.1002/gch2.202000038
10. Ding, W., Levine, R., Lin, C. and Xie, W. (2021). Corporate immunity to the COVID-19 pandemic. *Journal of Financial Economics*, 141(2), pp. 802-830. https://doi.org/10.1016/j.jfineco.2021.03.005
11. Hajian, N. and Hashemzadeh, M. (2024). The impact of COVID-19 outbreak on the investment of companies listed on the Tehran Stock Exchange. *Empirical Accounting Research Journal*, 14(2), pp. 213–240 (In Persian).
12. Jetter, M., Magnusson, L. M. and Roth, S. (2020). Becoming sensitive: Males' risk and time preferences after the 2008 financial crisis. *European Economic Review*, 128, A. 103512. https://doi.org/10.1016/j.euroecorev.2020.103512
13. Mankin, R. W., Jetter, E., Rohde, B. and Yasir, M. (2020). Performance of a low-cost acoustic insect detector system with Sitophilus oryzae (Coleoptera: Curculionidae) in stored grain and Tribolium castaneum (Coleoptera: Tenebrionidae) in flour. *Journal of Economic Entomology*, 113(6), pp. 3004-3010. https://doi.org/10.1093/jee/toaa203
14. Mwita, K. (2020). Effects of corona virus pandemic (COVID-19) on selected human resource management practices in Tanzania. *East African Journal of Social and Applied Sciences*, 2(2),

pp. 252-260.
15. Nasirzadeh, A., Jamalizadeh Noq, A. and Khoramnia. (2020). Investigating the status of stress, anxiety, depression, and resilience caused by the outbreak of COVID-19 in households in Anar city in 2020: A brief report. *Journal of Rafsanjan University of Medical Sciences*, 19(8), pp. 889-898.
16. Onwuegbuna, G. N., Nchuchuwe, F. F. and Ibrahim, H. K. (2021). Human resource management during COVID-19 pandemic: An insight on the challenges for human resource practitioners. *BVIMSR's Journal of Management Research*, 13(1), pp. 15-23.
17. Opatha, H. H. D. P. J. (2020). COVID-19 in Sri Lanka and seven HRM related recommendations against COVID-19. *Asian Journal of Social Sciences and Management Studies*, 7(2), pp. 152-157.
18. Opler, T., Pinkowitz, L., Stulz, R. and Williamson, R. (1999). The determinants and implications of corporate cash holdings. *Journal of Financial Economics*, 52(1), pp. 3-46. https://doi.org/10.1016/S0304-405X(99)00003-3
19. Pishbahar, E. and Rasouli, Z. (2019). Testing for neutrality and super-neutrality of money: Evidence from Iran's agricultural sector. *In Sustainable Agriculture and Agribusiness in Iran (pp. 151-161).* Singapore: Springer Singapore, Singapore. https://doi.org/10.1007/978-981-13-6283-5_10
20. Qin, X., Huang, G., Shen, H. and Fu, M. (2021). COVID-19 pandemic and firm-level cash holding moderating effect of goodwill and goodwill impairment. *In Research on pandemics (pp. 111-126).* Routledge, Abingdon, Oxfordshire.
21. Sabouri, N. (2022). Examining the role of financial flexibility in sustainable development on firm performance during the COVID-19 crisis considering tangible assets (Case study: Companies listed on Tehran Stock Exchange). *Journal of New Research Approaches in Management and Accounting*, 5(17), pp. 48-66 (In Persian).
22. Schweizer, L. and Nienhaus, A. (2017). Corporate distress and turnaround: integrating the literature and directing future research. *Business Research*, 10(1), pp. 3-47. https://doi.org/10.1007/s40685-016-0041-8
23. Shen, H., Fu, M., Pan, H., Yu, Z. and Chen, Y. (2021). The impact of the COVID-19 pandemic on firm performance. *In Research on Pandemics (pp. 81-98).* Routledge, Abingdon, Oxfordshire.
24. Shinde, P., Agraval, H., Srivastav, A. K., Yadav, U. C. and Kumar, U. (2020). Physico-chemical characterization of carvacrol loaded zein nanoparticles for enhanced anticancer activity and investigation of molecular interactions between them by molecular docking. *International Journal of Pharmaceutics*, 588, A. 119795. https://doi.org/10.1016/j.ijpharm.2020.119795
25. Zandi Gargari, M. and Amirhosseini, Z. (2016). The relationship between corporate governance guidelines and return on assets and earnings quality. *2nd International Conference on New Ideas in Management, Economics and Accounting*, Tehran, Iran (In Persian).
26. Zheng, C. and Zhang, J. (2021). The impact of COVID-19 on the efficiency of microfinance institutions. *International Review of Economics & Finance*, 71, pp. 407-423. https://doi.org/10.1016/j.iref.2020.09.016
27. Zheng, M. (2022). Is cash the panacea of the COVID-19 pandemic: Evidence from corporate performance. *Finance Research Letters*, 45(1), A. 102151. https://doi.org/10.1016/j.frl.2021.102151

IJAAF; Vol. 10, No. 2, Spring 2026, Serial 35

Iranian Journal of
Accounting, Auditing & Finance

Ferdowsi University of Mashhad

Quarterly

RESEARCH ARTICLE

The Analysis of Peer Effects in Environmental, Social, and Governance Disclosure: The Moderating Role of Firm Experience and Size

Hamed Parvizikia, Daruosh Foroghi*, Narges Hamidian, Hasan Fattahi Nafchi

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

How to cite this article:
Parvizikia, H. , Foroghi, D. , Hamidian, N. and Fattahi Nafchi, H. (2026). The Analysis of Peer Effects in Environmental, Social, and Governance (ESG) Disclosure: The Moderating Role of Firm Experience and Size. Iranian Journal of Accounting, Auditing and Finance, 10(2), 37-54. doi: 10.22067/ijaaf.2026.47506.1530
https://ijaaf.um.ac.ir/article_47506.html

ARTICLE INFO

Article History
Received: 2025-09-01
Accepted: 2025-11-18
Published online: 2026-04-12

Abstract

This study examined the peer effect on Environmental, Social, and Governance (ESG) disclosure in industries through imitative and reciprocal mechanisms and assessed the moderating roles of experience and company size. From 2012–2024, 117 different firms were listed on the Tehran Stock Exchange, making them part of the statistical sample. Principal Component Analysis (PCA) was used to generate the composite ESG index, and linear regression models were used to evaluate the hypotheses. A company's ESG disclosure was positively affected by its peers' ESG practices, according to the results, which showed a substantial intra-industry peer effect on ESG disclosure. The intra-industry peer impact was unexpectedly amplified by company experience. Large companies operating in the same industry showed mimetic peer influence, whereas small and large companies did not. On the other hand, whereas large and small enterprises were found to have a mimetic peer impact, small firms did not exhibit any reciprocal peer effect. This study took a fresh look at the impact of firms' size and experience as moderators of intra-industry peer effects on environmental, social, and governance (ESG) disclosure in Iran, differentiating between reciprocal and imitative dynamics. Policymakers and practitioners can use these findings to improve corporate transparency, and the literature on peer influence in ESG disclosure benefits from them as well.

Keywords:
ESG, Experience, Firm Size, Peer Effect

https://doi.org/10.22067/ijaaf.2026.47506.1530

CC BY

NUMBER OF REFERENCES
38

NUMBER OF FIGURES
-

NUMBER OF TABLES
5

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

*Corresponding Author: Daruosh Foroghi
Email: foroghi@ase.ui.ac.ir
Tel: 09131163291
ORCID: 0000-0002-7164-6728

1. Introduction

In recent years, investing in corporate environmental, social, and governance (ESG) information disclosure has been considered of significant importance. The disclosure of ESG information enhances a company's non-financial performance indicators, such as consumer satisfaction and market acceptance, which signals stakeholders that the organization is ethically committed to environmental sustainability (Chen and Ma, 2017). As a result of this dedication, the company's standing with its creditors improves. Capelle-Blancard and Petit (2019) found that investors can react negatively to inconsistent ESG disclosure procedures.

Companies' actions and decisions within the same industry affect one another, a phenomenon known as the intra-industry peer effect, a prominent feature of corporate ESG disclosure. Even in Iran's stock market, companies' ESG reporting procedures are affected by their peers, whether they realize it or not. There are mainly two ways in which the peer effect shows itself: logical imitation and reciprocal imitation. When it comes to logical imitation, smaller businesses often mimic larger ones to stay competitive, secure cheap financing, and keep operations running smoothly. In reciprocal imitation, firms of similar size, whether large or small, tend to mimic each other's practices (Hu et al., 2022). Consequently, firm size is expected to significantly shape the intra-industry peer effect in ESG disclosure.

In addition to firm size, firm-specific factors, such as experience, also affect the peer effect on ESG information disclosure. Firm experience can moderate the influence of peers on firms (Leary and Roberts, 2014). Specifically, firms with limited ESG reporting experience are prone to imitate their more experienced counterparts to produce comprehensive reports and enhance their chances of survival in the market (Chen and Ma, 2017). Accordingly, greater experience is expected to reduce a firm's tendency to imitate its peers in ESG disclosure. In other words, more experienced firms are less likely to be influenced by their intra-industry peers' ESG reporting practices.

Peer effects in this domain, especially in emerging markets such as Iran, remain underexplored despite growing attention to ESG disclosures. Little is known about ESG-specific imitation, even though firms often mimic peers in financial decisions (Leary and Roberts, 2014; Adhikari and Agarwal, 2018). This study investigates the moderating role of firm size and experience in ESG disclosure behavior. Larger firms often serve as benchmarks (Haunschild and Miner, 1997; Hu et al., 2022), but the responses of smaller or younger firms are unknown. Prior research has mainly focused on financial outcomes, such as earnings management (Kordestani and Jafari Souq, 2023; Golijani and Jafar Salehi, 2024), overlooking strategic ESG dynamics. This research provides deeper insight into the effect of peer dynamics on ESG transparency across industries, filling a significant gap in the Iranian capital market literature by distinguishing between rational and reciprocal imitation (DiMaggio and Powell, 2022). Thus, this study seeks to evaluate the peer effect on ESG disclosure, with firm size and experience as moderating factors. For managers, the findings provide strategic insights into leveraging ESG disclosure not only for regulatory compliance but also for competitive positioning by better understanding and anticipating peer behavior, enabling more effective resource allocation, reputation management, and long-term value creation. Imitation helps policymakers better understand potential regulatory dynamics, leading to better regulation design that promotes transparency and reduces greenwashing, particularly in emerging markets where formal ESG standards are still maturing.

This paper first reviews the theoretical and empirical foundations, then presents the methodology and findings. Finally, it discusses conclusions, practical implications, and directions for future research.

2. Literature review

The peer effect, originating in sociological theory, describes the tendency for individuals to adjust their behavior in response to others' actions, often leading to patterns that reflect broader group dynamics (Manski, 1993), which is not only true for people, but also for entities such as companies or enterprises whose behavior and decision-making can also be greatly influenced by the actions of others in their environment. Sociological theories emphasize that imitation based on social ties is a widespread phenomenon and that there are strong similarities in firm behavior.

Peer firms with structural similarities in industry and technology can significantly affect one another's performance across various domains. Academic studies have paid little attention to the phenomenon of financial and industrial strategies converging through mutual influence (Duong et al., 2015). This has been further supported by empirical research showing that companies often mimic their competitors' ESG policies and disclosures (Adhikari and Agarwal, 2018; Leary and Roberts, 2014). Such imitative behavior is well-documented in the corporate context, with evidence suggesting that firms often adopt and replicate information from one another (Yang et al., 2017).

In recent years, rapid industrial expansion and the resulting environmental degradation have intensified global concern regarding sustainability, placing increasing pressure on firms to adopt responsible ESG practices. Simultaneously, stakeholders are calling for increased corporate accountability, especially after several corporations have been embroiled in scandals that point to governance lapses that not only cause direct parties to suffer economic losses but also affect many innocent bystanders. This is particularly a source of legitimacy and acceptance for firms in developing economies pursuing deeper engagement with global markets, where this position on ESG has emerged as a clear imperative. Corporate sustainability is commonly understood as a holistic approach, conducted within an organization's portfolio of activities to meet current stakeholder needs without compromising resources for future generations. Although the concept has been widely discussed in the literature, its precise definition remains debated (Chen and Ma, 2017). The ESG framework is a widely accepted model for evaluating non-financial performance that encompasses three key dimensions: environmental, social, and governance.

Environmental: This dimension varies across industries. For instance, carbon emissions are generally more essential for manufacturing firms than financial institutions (Eccles and Serafeim, 2013). Common environmental concerns include climate risks, pollution, energy and water management, and biodiversity loss. As sustainability becomes an increasingly important business priority, these concerns are increasingly shaping firm strategies.

Social: The social piece includes human rights, labor conditions, employee welfare, and consumer relations, each of which is important in different situations. Labor practices in developing countries are comparatively lower-cost, and thus their human rights incidents are more pronounced (Eccles and Serafeim, 2013). Tackling these challenges is essential for upholding stakeholder confidence and purposeful accountability.

Governance: Governance encompasses systems and procedures that direct corporate behavior and decision-making, which concerns the dynamic among management, shareholders, and other stakeholders (Bernardi and Stark, 2018). Despite this, the recent literature has begun to examine how firms respond to the governance practices of other firms in their industry, demonstrating a growing recognition of peer effects.

Emerging ESG tensions compel firms to increasingly refer to peer behavior when determining their disclosure approach, especially in complex or uncertain situations. Two theoretical frameworks are used to explain this tendency: information-based theory and competition-based theory. The information-based theory (Lieberman and Asaba, 2006) posits that firms copy one another largely due to incomplete information. Without clear standards to work from, firms watch and copy what peers do to mitigate uncertainty and help them decide how to behave. Under this theoretical lens, social learning theory (Bikhchandani et al., 1998) has been widely adopted to highlight that imitation of leading firms is generally undertaken by fewer resourceful or experienced firms, aiming to achieve conformity and legitimacy with best practices (Dodgson, 1993). These top firms, endowed with deeper pockets, strategic foresight, and visibility, are also information conduits that others follow suit (Wang et al., 2024). Such a peer effect is stronger in ESG-related domains, where follower firms can learn from their peers' disclosures and strategies with less risk of externalities (Li et al., 2018). This dynamic has empirical backing; for example, Li et al. (2018) and Zhi-guo and Bo-yi (2020) found that firms mimic financial behavior, such as capital structures and dividend policies, based on regional or local peers, implying that peer influence is not limited to ESG decisions but extends to broader strategic choices.

The competition-based theory adds a strategic rivalry perspective to the information-based view. In competitive markets, companies imitate their peers not just to reduce uncertainty, but also to avoid being outperformed. For example, companies employ similar investment and sustainability strategies to offset periodic costs and prevent strategic flops (Xi et al., 2024). Dynamic competition theory reinforces this notion, suggesting that firms adjust ESG disclosures to maintain legitimacy and competitive edge (Wan et al., 2016). As ESG performance becomes a benchmark in investor decisions and regulatory evaluations, the pressure to emulate increases (Kaustia and Rantala, 2015). Firms closely monitor competitors' ESG-specific activities to maintain market relevance (Gimeno et al., 2005). Delay in keeping pace with peers can affect stakeholder confidence, while benchmarking helps get ESG efforts back on track and improves visibility. Gu and Shen (2024) investigate peer effects on environmental information disclosure in Chinese heavy-polluting industries. Their findings show that peer effects shape both "real green" and "fake green" behaviors. True pro-environmental behaviors are largely competitive in nature, but fake green or symbolic actions, we find, are more indirect and akin to social learning. In addition, Hu et al. (2022) report an intra-industry peer effect in corporate environmental information disclosure.

Such convergence, driven partly by information gaps and marketplace pressures, makes the widespread dissemination of ESG practices possible. When leading firms launch high-profile ESG initiatives, this often triggers imitative behavior (Grewal et al., 2021), prompting other companies to follow suit; in turn, this accelerates the emergence of standardized ESG disclosures across industries. From the theoretical arguments explained above, we propose our first hypothesis listed below:

H1: There is a significant intra-industry peer effect in ESG disclosure.

Research has shown that anchoring managerial capabilities in experiential knowledge is a core element of their development. Accumulated professional experience largely reflects an individual's expertise (Bonaiuto et al., 2003), especially the ability to assume important responsibilities or to assess a firm's position within an industry. Ultimately, learning from diverse experiences enhances both strategic judgment and operational efficiency (and firms led by more experienced managers outperform those guided by less experienced managers; managers, Demerjian et al., 2012). Managerial experience for corporate governance is essential if we consider the core functions of management, like planning, decision-making, etc., by which it keeps the managerial actions aligned with strategic goals. Also, wiser and more experienced minds are usually better equipped to handle complex decisions, which adds to organizational maturity. Further, given the competitive landscape today, ESG disclosure has evolved into a significant means of shaping stakeholder perceptions and sustaining a firm's market position, thereby enhancing organizational success (Hambrick and Mason, 1984).

The level of experience accumulated by a firm, reflected in its resources, knowledge accumulation, and organizational maturity, is an integral factor influencing how much it relies on peer behavior when making ESG disclosure decisions. Theoretical and empirical studies favor this perspective. However, firms that have less experience or fewer resources rely heavily on their peers' behavior due to heightened information uncertainty (Lieberman and Asaba, 2006). These firms rely on leading companies to reduce uncertainty and inform their behavior by observing the actions of more-experienced firms. In the same vein, social learning theory (Bikhchandani et al., 1998) suggests that firms with less experience often adopt the practices of high-profile peers to obtain legitimacy and align with what is perceived as best practice. In the absence of internal knowledge or well-developed ESG norms, firms tend to rely significantly more on disclosure and strategies implemented by other organizations (Dodgson, 1993). By contrast, seasoned firms, which tend to have more robust resources and deeper strategic awareness, are in a stronger position not only to produce their own information but also to rely less on imitation, often serving as reference firms for others (Wang et al., 2024).

In fact, competition-based theory (Wan et al., 2016) supports the aforementioned argument by indicating that a firm's level of experience further influences its response to competitive pressure. Less seasoned firms are more inclined to mimic their competitors' ESG strategies to avoid falling in competitive markets. They also try to reduce repeat expenses and mitigate the risk of strategic errors by emulating these behaviors (Hu et al. 2022). On the contrary, more established companies are usually far better suited to innovate and build their own ESG strategies. Thus, they are less swayed by peers' behavior, even as they closely follow and respond to competitors to secure their competitive standing (Gimeno et al., 2005).

Based on theoretical predictions, we provide robust empirical evidence that firms' tendency to copy peers in ESG disclosure declines with increasing firm-level experience. For example, Li et al. (2018) demonstrated that firms with less experience and fewer resources in financial matters (e.g., capital structure decisions) are more influenced by the behavior of local or regional peers, from whom they receive phylogenetic and learning benefits, and this is also true for ESG disclosure practices. Likewise, Zhi-guo and Bo-yi (2020) indicate that in emerging markets, less experienced firms frequently imitate the financial and sustainability policies of several leading companies to adopt market practice. Under institutional and market pressures, particularly regarding ESG-related issues, Kaustia and Rantala (2015) demonstrate that nascent firms imitate their peers to maintain legitimacy in the eyes of investors and other stakeholders. This happens particularly in sectors where ESG standards are not yet well established.

By contrast, firms with greater experience, supported by stronger internal systems and well-established reputations, are generally less reactive to such pressures. Nevertheless, they may still align certain aspects of their ESG disclosures with those of their peers to preserve their competitive position (Hu et al., 2022; Xi et al., 2024). In addition, Hu et al. (2022) show that firm experience plays a significant moderating role in the peer effect associated with environmental information disclosure, an important component of ESG reporting. Their findings suggest that firms with limited experience, often lacking sufficient knowledge, resources, or internal infrastructure, rely more heavily on peer imitation.

Overall, these results show that experience is a relevant moderator of the ESG disclosure peer effect. Relatively inexperienced firms are more inclined to imitate because they have limited resources and knowledge. In contrast, highly experienced firms tend to act as informational reference points and pursue less imitative, novel strategies. Based on the theoretical arguments and empirical evidence discussed above, the second hypothesis is proposed as follows:

H2: A company's experience reduces the influence of intra-industry peers on its ESG disclosures.

Firm size plays a substantial role in corporate governance and ESG disclosure by shaping stakeholder perceptions and practices guiding disclosure; establishing the framework of a firm's social responsibility initiatives, and the extent to which they will engage with industry peers. Larger firms frequently act as reference points within their industries, encouraging smaller or less established companies to follow their lead (Haunschild and Miner, 1997). As a result, strategic choices—including ESG reporting, often become similar across firms due to external pressures and peer influence (Leary and Roberts, 2014).

The size of the firm thus significantly mediates peer influences on ESG disclosure through both information and competition channels. According to information-based theory, companies with incomplete or asymmetric information are more likely to follow the actions of their peers that they perceive as knowledgeable or competent (Lieberman and Asaba, 2006). Relatedly, social learning theory indicates that in environments with few prominent ESG benchmarks, large organizations are more likely to observe their better-informed peers and emulate their practices to attenuate uncertainty and buttress legitimacy (Bikhchandani et al., 1998). In practice, resource-rich large firms with greater market visibility and stronger reporting capabilities have been the ones informing the market about ESG best practices. These signals prompt other smaller firms to follow suit with similar disclosure strategies. Such mimicry allows smaller organizations to reduce decision-making risk and builds confidence amongst stakeholders (Li et al., 2018; Wang et al., 2024).

Concomitantly, as the competition-based theory indicates that firms operating in competitive markets establish their strategies relative to those of their competitors, they follow this path when ESG performance is seen as relevant by stakeholders and regulators (Xi et al., 2024; Wan et al., 2016). Within this logic, peer effects can arise via two channels. In large firms, companies often mimic competitors and improve their ESG practices in line with competitors' target-driven behavior (Haunschild and Miner, 1997; Hu et al., 2022). Simultaneously, firms in the same niche and of similar size often follow one another to avoid falling into a negative equilibrium (Hu et al., 2022). Such back-and-forths tend to be most pronounced within smaller firms, as their analogous constraints and exposure to aligned market pressures compel them to surveil for and mimic the ESG strategies of peer companies.

These trends are well-documented in empirical research. Larger A-share firms have more ESG disclosure, according to Zhang and You (2024), while Huang et al. (2024) find that firms with information disadvantages often mimic the ESG practices of industry leaders. In addition, Hu et al. (2022) provide direct evidence of intra-industry spillover effects in environmental disclosure. Overall, these findings imply that firm size plays a role not only in shaping the aggregate level of ESG disclosure but also in the peer effect mechanism by which firms with peers share information about sustainability, leading to faster convergence in transparency outcomes within an industry. On this basis, the following hypotheses are advanced:

H3*:* Large firms exhibit a reciprocal intra-industry peer effect in ESG disclosures with other large firms.

H4*:* Large firms do not exhibit an intra-industry peer effect in ESG disclosures with small firms.

H5*:* Small firms demonstrate an intra-industry peer imitation effect in ESG disclosures with large firms.

H6*:* Small firms exhibit a reciprocal intra-industry peer effect in ESG disclosures with other small firms.

The present study introduces several innovative dimensions which distinguish it from prior research on peer effects and ESG disclosures. Although peer effects have been extensively investigated in the social sciences, particularly in sociology, their application in accounting research, especially within the Iranian context, remains limited. While Joodaki et al. While Golijani and Jafar Salehi (2024) explore the influence of environmental disclosure on earnings management, and Kordestani and Jafari Souq (2023) examine these relationships with respect to capital market pressures to achieve profits, none provide a substantive examination of the motivations underlying positive investment outcomes. As a result, the dynamics underlying ESG disclosure, and how peers influence them, remain insufficiently explored.

This study fills that gap in four important ways. First, it provides the first empirical analysis of intra-industry peer effects on ESG disclosure within Iran's capital market, using panel data from firms listed on the Tehran Stock Exchange. Second, by separating imitative peer effects (small firms imitating large firms) from reciprocal peer effects (mutual influence among peers), the study identifies behavioral patterns that had not yet been documented in research within the domestic context. Third, it demonstrates that firm size and managerial experience jointly moderate these peer effects, adding more nuance to the understanding of how organizational characteristics drive ESG disclosure behavior. Finally, the study devises a context-specific theoretical framework that can inform practitioners and policymakers in enacting more transparent and effective ESG reporting through blending social learning and competition–based theories with a rigorous understanding of Iran's regulatory and cultural backdrop.

These four innovations have obvious practical advantages. First, substantive evidence on intra-industry peer effects enables regulators to design disclosure guidelines tailored to firm size and industry characteristics. Second, identifying the presence of imitative and reciprocal peer effects enables managers to tailor their benchmarking strategy, either to mimic industry leaders or to compare with peers of a similar size. Third, recognizing how the size and experience of firms affect dynamics among contemporaries enables investors to make more informed decisions about the risks and opportunities of ESG disclosure. Finally, a context-specific framework that integrates social learning and competition theories helps policymakers design measures that promote transparent and effective ESG reporting across firms with different characteristics.

3. Research methodology

To test the hypotheses, linear regression models were estimated using EViews 12. The study population consists of firms listed on the Tehran Stock Exchange over the period 2012–2024. The sample was selected based on several criteria, including that firms must have been listed on the TSE prior to the study period; their fiscal year had to end in March to ensure comparability; they could not belong to the insurance, banking, brokerage, or investment sectors; and the necessary annual data had to be available for calculating the study variables. Firms that changed their fiscal year during the sample period were also excluded. In addition, because some variables are measured at the industry level, each industry had to include at least 5 firms. Applying these criteria yielded a final sample of 117 firms across 12 industries listed on the Tehran Stock Exchange.

3.1. Research models

The hypotheses were tested following the methodology of Hu et al. (2022). The regression model is as outlined in Model 1 for testing the first hypothesis.

$$ESG_{i,t}=\alpha+\beta_1 ESG\_Industry_{j,t-1}+\beta_2 Size_{i,t}+\beta_3 CR_{i,t}+\beta_4 LEV_{i,t}+\beta_5 ROA_{i,t}+\beta_6 QTobin_{i,t}+\beta_7 Grow_{i,t}+\beta_8 Auditor\_size_{i,t}+\beta_9 Opin_{i,t}+\varepsilon_i \quad (1)$$

$ESG_{i,t}$: The Environmental, Social, and Governance (ESG) disclosure score of company i in year t.
$ESG_{Industry\ j,\ t-1}$: The average ESG disclosure score for industry j (excluding company i) in year t−1

SIZE: Company size
CR: Concentration Ratio
LEV: Financial leverage
ROA: Return on Assets
Q_{Tobin}: The market value ratio of assets to the book value of the company's assets.
Grow: Capture of the operating income growth rate
$Auditor_{Size}$: The size of the firm's auditor
Opin: The audit opinion

The coefficient β_1 captures the effect of industry-level ESG disclosure on firm-level ESG disclosure. In line with the first hypothesis, β_1 is expected to be positive and statistically significant, indicating the presence of peer effects in ESG disclosure within industries. To test the second hypothesis, the regression specification presented in Model 2 is employed.

$$ESG_{i,t} = \alpha + \beta_1 ESG\_Industry_{j,t-1} + \beta_2 AGE_{i,t} + \beta_3 ESG\_Industry_{j,t-1} \times AGE_{i,t} + \beta_4 Size_{i,} + \beta_5 CR_{i,t} + \beta_6 LEV_{i,t} + \beta_7 ROA_{i,t} + \beta_8 Qtobin_{i,t} + \beta_9 Grow_{i,t} + \beta_{10} Audi\_Size_{i,t} + \beta_{11} Opin_{i,t} + \varepsilon_{i,t} \quad (2)$$

AGE: The firm's experience, along with the remaining variables, is consistent with the specifications outlined in Model 1. Under this hypothesis, the coefficient β3 is expected to be negative and significant, in which case, the hypothesis is not rejected.

To test the third and fourth hypotheses, firms are categorized into small and large companies. The size variable is split at the median; firms with sample median values are coded as large, and those below the median are categorized as small. To test these two hypotheses, the regression relationship in Model 3 is estimated for the large company sample.

$$ESG_{i,t} = \alpha + \beta_1 B\ ESG\_Industry_{j,t-1} + \beta_2 S\ ESG\_Industry_{j,t-1} + \beta_3 CR_{i,t} + \beta_4 LEV_{i,t} + \beta_5 ROA_{i,t} + \beta_6 QTobin_{i,t} + \beta_7 Grow_{i,t} + \beta_8 Auditor\_Size_{i,t} + \beta_9 Opin_{i,t} + \varepsilon_i \quad (3)$$

B $\mathbf{ESG_{Industry j,\mathit{t}-1}}$: The environmental, social, and governance (ESG) disclosure score for large firms within industry *j*, excluding the score of company *i*.

S $\mathbf{ESG_{Industry\ j,\ t-1}}$: The ESG disclosure score for small companies within industry j (excluding company i) is constructed, while the remaining variables are consistent with those specified in Model 1.

Based on the research hypothesis, large companies may imitate the ESG disclosure practices of other large companies within the same industry. However, no such imitation occurs with small companies in the industry. The coefficient β_1 captures the reciprocal peer effect of ESG information disclosure among large companies within the same industry. The coefficient β_1 captures the reciprocal peer effect of ESG information disclosure among large companies within the same industry. When β_1 is positive and statistically significant, the third hypothesis cannot be rejected. Conversely, the coefficient β_2 measures the peer effect of ESG information disclosure between large companies and small companies within the same industry. When β_2 is statistically insignificant, the fourth hypothesis is accepted.

The regression relationship in Model 3 is also estimated for the sample of small companies to test the fifth and sixth hypotheses. Small companies are expected to imitate the ESG disclosures of large companies in their industry and also engage in reciprocal information imitation with small companies in the same industry. In other words, the coefficient β1 captures the peer-imitation effect of ESG disclosures across small and large companies within the same industry. When this coefficient is positive and significant, the fifth hypothesis is accepted. The coefficient β2 captures the reciprocal peer effect of ESG disclosures between small companies in the same industry. When this coefficient is positive and significant, the sixth hypothesis will not be rejected.

3.2. Research variables

3.2.1. Dependent variable

Following Fakhari et al., a weighted scoring model is used to compute the ESG disclosure score for companies listed on the Tehran Stock Exchange (2017). An adapted checklist of environmental (e.g., carbon emissions, waste management), social (e.g., employee welfare; community engagement), and governance (e.g., board independence; transparency) indicators based on Fakhari et al., which helps in the analysis of board reporting) (Madiraju et al. , 2022). A score of 1 is assigned if each indicator is present, and a score of 0 if absent. ESG scores are calculated separately for the environmental (E), social (S), and governance (G) dimensions. Principal Component Analysis (PCA) is then employed to obtain the weights for these dimensions from the research data to create the composite ESG index. The ESG score is obtained using Equation 4:

$$ESG = 0.5969*E + 0.6039*S + 0.5280*G \qquad (4)$$

These coefficients are derived from the PCA output and reflect the relative importance of each dimension. Here, ESG represents the composite Environmental, Social, and Governance reporting score, and E, S, and G denote the respective dimension scores. PCA is used exclusively to weight the ESG index rather than for other analyses.

3.2.2. Independent variable

The independent variable is the disclosure of environmental, social, and governance reporting at the industry level ($ESG_{Industry}$). The ESG disclosure score of the companies within each industry (excluding company i in the same industry) is averaged to calculate this variable.

3.2.3. Moderator variables

The studied moderating variables, following Hu et al. (2022), are company size and experience, which are explained as follows:

Size*:* The natural logarithm of total assets is used to calculate this variable.

Company Experience (Age)*:* The company experience is measured using the company's age. The company age is calculated by taking the natural logarithm of the difference between the number of years since the company's listing date on the stock exchange and the current year.

3.2.4 Control variables

In this study, the variables influencing the ESG disclosure of companies were controlled, following the research of Hu et al. (2022), Khozein et al. (2018), and Akbas (2014). These variables are as follows:

Company Size (SIZE): The natural logarithm of total assets, which is used as a control variable in hypotheses 1 and 2, and as a moderating variable in hypotheses 3 to 6.

Ownership Concentration (CR): Equal to the sum of the ownership percentage held by shareholders owning at least 5% of the company's shares (Bahramoghadam et al., 2013).

Financial Leverage (LEV): The ratio of total liabilities to total assets.

Return on Assets (ROA): The ratio of net profit to total assets.

Q_{Tobin}: The ratio of the market value of assets (book value of debt plus the market value of equity) to the book value of the company's assets.

Operating Income Growth Rate (Grow): The difference between the operating income of year t and year t-1 divided by the operating income of year t-1.

Auditor Size ($Auditor_{size}$): If a firm is audited by the Audit Organization or an auditing firm with an "A" rating, the variable is assigned a value of 1; otherwise, it is assigned a value of 0.

Audit Opinion (Opin): If the auditor has issued an adverse or qualified opinion, the value is 1; otherwise, it is 0.

4. Results

Table 1 presents the descriptive statistics for the key variables under study.

Table 1. Descriptive statistics

Variable	Mean	Median	Max	Min	Std. Dev
ESG	9.294	9.501	19.705	0.000	3.787
$ESG_{Industry}$	8.927	8.844	15.388	4.275	1.774
SIZE	15.220	14.927	21.899	10.532	1.858
AGE	2.871	2.944	4.025	0.693	0.462
CR	0.683	0.720	1.000	0.000	0.208
LEV	0.536	0.529	2.327	0.017	0.230
ROA	0.165	0.140	1.140	-0.400	0.163
QTobin	2.561	1.899	32.702	0.490	2.490
GROW	0.609	0.297	23.725	-15.786	2.557
$Auditor_{size}$	0.674	1.000	1.000	0.000	0.468
Opin	0.179	0.000	1.000	0.000	0.384

For the dependent variable—the ESG disclosure score—the mean of 9.29 indicates that most observations are clustered around this level. Similarly, the independent variable—industry-level ESG disclosure—has a mean value of 8.92, indicating that its observations also tend to cluster around this level. The company-level ESG disclosure reaches a maximum of 19.70, with a mean of 9.29, while industry-level ESG disclosure reaches a maximum of 15.38 (mean= 8.92). This difference also implies a greater level of ESG disclosure at the individual firm level compared to the aggregated (industry) level, indicating how important firms view transparency around ESG. The ownership concentration variable (min = 0, max = 100, mean = 0.68) indicates that many companies are majority-owned by large shareholders. Tobin's Q (mean = 2.65) indicates that the stocks of the studied companies have relatively high risk. In the case of multiple firms, we have some level of growth. In others, a negative growth variable, which is explained by its average fee of up to 0.606 and a minimum value of up to -15.7 for this examine revenue increase rate. The average of size and auditor's opinion variables have been recorded as 0.67 and 0.18, respectively; this demonstrates that most companies have gone through acceptable audit processes. For the dependent variable, the standard deviation is 3.78, indicating that these firms' ESG disclosure scores differ from the average by about 3.78 points. The maximum and minimum parameters indicate the dispersion of observations, i.e., the range of values. Moreover, the min-max values for the dependent and independent variables, i.e., 19.70 and 15.38, respectively, provide clarity regarding the variability present within the data itself.

4.1. Hypothesis testing

The regression models in the study are estimated while controlling for year and industry effects. The White test was used to assess heteroskedasticity, and the Breusch-Godfrey test was used to assess first-order autocorrelation. The p-values for both tests were below 5%, indicating heteroskedasticity and autocorrelation. The Newey-West correction method was employed to mitigate these effects in the regression estimations. Additionally, multicollinearity among the explanatory variables was examined using the Variance Inflation Factor (VIF) statistic. The VIF values for all explanatory variables in the models were below 10, indicating no significant multicollinearity issues.

4.1.1. Test of H1

Hypothesis 1 predicts an intra-industry peer effect in ESG disclosure. Table 2 summarizes the Regression Model 1 results.

Table 2. Regression results for H1

$ESG_{i,t} = \alpha + \beta_1 ESG_{Industryj,t-1} + \beta_2 Size_{i,t} + \beta_3 CR_{i,t} + \beta_4 LEV_{i,t} + \beta_5 ROA_{i,t} + \beta_5 ROA_{i,t} + B_6 QTobin_{i,t} + \beta_7 Grow_{i,t} + \beta_8 Auditor_{Sizei,t} + \beta_9 Opin_{i,t} + \varepsilon_i$

Variable	Coefficient	Std. Error	T-Statistic	P-Value
$ESG_{Industry}$	0.164	0.079	2.072	0.038
SIZE	0.566	0.070	8.040	0.000
CR	0.571	0.462	1.234	0.217
LEVRAGE	2.204	0.778	2.832	0.004
ROA	3.920	0.156	3.389	0.000
Q_{TOBIN}	-0.137	0.045	-3.429	0.002
GROW	0.004	0.043	0.106	0.915
$Auditor_{Size}$	0.909	0.242	0.375	0.707
OPIN	0.319	0.272	1.172	0.241
C	-2.267	1.236	-1.834	0.066
Year & Industry Effect	yes			
Adjusted R^2	0.191			
F-Statistic (Prob)	31.210 (0.000)			

The F-statistic is 2100.31, with a significance level of 0.0000, indicating that the overall regression model is statistically significant. The adjusted R-squared value is 0.1931, suggesting that approximately 19% of the variations in the dependent variable are explained by the independent and control variables. The significance level for the industry-level ESG disclosure score is 0.0384, which is below the 5% error threshold. Additionally, the coefficient for this variable is 0.1643, indicating that peer ESG disclosure has a positive and significant impact on a company's ESG disclosure. In other words, there is an intra-industry peer effect in ESG disclosure. Therefore, the first research hypothesis is not rejected at the 95% confidence level, indicating that industry-level ESG disclosure strengthens firm-level ESG disclosure. The coefficient for this variable, as expected, is positive at 0.1643, indicating that a 1% change in the independent variable (industry-level ESG disclosure) results in a 0.1643% change in the dependent variable (firm-level ESG disclosure score).

4.1.2. Test of H2

Hypothesis 2 expresses that firm experience attenuates the intra-industry peer effect in ESG disclosure. Table 3 presents the results from Regression Model 2, which tests this hypothesis.

Table 3. Regression results for H2

$ESG_{i,t} = \alpha + \beta_1 ESG\_Industry_{j,t-1} + \beta_2 AGE_{i,t} + \beta_3 ESG\_Industry_{j,t-1} \times AGE_{i,t} + \beta_4 Size_{i,} + \beta_5 CR_{i,t} + \beta_6 LEV_{i,t} + \beta_7 ROA_{i,t} + B_8 Qtobin_{i,t} + \beta_9 Grow_{i,t} + \beta_{10} Audi\_Size_{i,t} + \beta_{11} Opin_{i,t} + \varepsilon_{i,t}$

Variable	Coefficient	Std. Error	T-Statistic	P-Value
$ESG_{Industry}$	-0.638	0.299	-2.135	0.032
AGE	-0.834	0.973	-0.857	0.391
$ESG_{Industry*AGE}$	0.249	0.103	2.402	0.016

SIZE	0.475	0.063	7.436	0.000
CR	-0.138	0.439	-0.315	0.752
LEVRAGE	1.109	0.508	2.182	0.029
ROA	3.773	0.824	4.575	0.000
Q TOBIN	-0.146	0.047	-3.102	0.002
GROW	-0.015	0.035	-0.437	0.661
$Auditor_{Size}$	-0.124	0.202	-0.612	0.540
OPIN	0.773	0.232	3.319	0.000
C	3.176	3.010	1.026	0.304
Year & Industry Effect		yes		
Adjusted R^2		0.188		
F-Statistic (Prob)		25.675 (0.000)		

According to the second hypothesis, the coefficient for the variable $ESG_{Industry}$×AGE was expected to be negative and significant. However, as shown in Table 3, the coefficient is significant at the 95% confidence level (p=0.0164), but contrary to expectations, it is positive. Therefore, the second hypothesis is rejected at the 95% confidence level. Since the coefficient is positive (0.2495), a company's experience enhances the intra-industry peer effect on ESG disclosure. The intra-industry peer effect on ESG disclosure increases as a company matures.

The positive coefficient suggests that firm experience strengthens the industry-level peer effect on ESG disclosure. In particular, the peer effect on a firm's ESG disclosure practices strengthens with experience. This conclusion is consistent with information-based theory, which argues that firms experience informational gaps in their reporting and seek to adopt peers' disclosure practices, thereby improving the quality and coverage of their own ESG information. Consequently, more mature firms have greater impetus and capacity to absorb and incorporate ESG-related information arising from industry peers, thereby reinforcing the existing peer influence at the industry level. This finding highlights the influence of experience on firms' strategic responses to information deficits and on their dependence on peer-group comparisons in response to ESG disclosure practices.

4.1.3. Testing of the H3, H4

To test the third and fourth hypotheses, Regression Model (3) has been estimated for the sample of large firms. A summary of the results from estimating this model for these two hypotheses is presented in Table 4.

According to the third hypothesis, large firms are expected to exhibit a reciprocal intra-industry peer effect in ESG disclosure with other large firms. The significance level for the variable representing ESG disclosure for large firms relative to other large firms in the same industry is 0.0177, which is below the 5% error threshold. Additionally, the coefficient for this variable is 0.1438, indicating that ESG disclosure by large firms within an industry has a positive and significant impact on the ESG disclosure of other large firms. In other words, there is a reciprocal intra-industry peer effect in ESG disclosure among large firms in the same industry, meaning that large firms not only rely on their own information but also mutually imitate each other's ESG disclosure practices. Therefore, the third hypothesis is not rejected at the 95% confidence level.

Table 4. Regression results for H3, H4

$ESG_{i,t} = \alpha + \beta_1 B\ ESG\_Industry_{j,t-1} + \beta_2 S\ ESG\_Industry_{j,t-1} + \beta_3 CR_{i,t} + \beta_4 LEV_{i,t} + \beta_5 ROA_{i,t} + B_6 QTobin_{i,t} + \beta_7 Grow_{i,t} + \beta_8 Auditor\_Size_{i,t} + \beta_9 Opin_{i,t} + \varepsilon_i$

Variable	Coefficient	Std. Error	T-Statistic	P-Value
B $ESG_{Industry}$	0.143	0.060	2.378	0.017
S $ESG_{Industry}$	0.079	0.072	1.096	0.273
CR	-0.052	0.631	-0.083	0.933
LEVRAGE	1.114	0.794	1.402	0.161

ROA	4.425	1.402	3.155	0.001
Q TOBIN	-0.158	0.088	-1.787	0.074
GROW	0.077	0.030	2.579	0.010
Auditor_ Size	0.622	0.309	2.013	0.044
OPIN	0.924	0.307	3.005	0.002
C	5.931	0.892	6.644	0.000
Year & Industry Effect	yes			
Adjusted R^2	0.129			
F-Statistic (Prob)	10.428 (0.000)			

According to the fourth hypothesis, large firms are not expected to exhibit a peer effect in ESG disclosure through imitation with small firms within the same industry. As shown in the results in Table 4, as expected, the significance level for the variable representing ESG disclosure by large firms with small firms in the same industry is 0.2731, which is above the 5% error threshold, indicating that there is no peer effect through imitation in ESG disclosure between large and small firms within the same industry. In other words, ESG disclosure by small firms does not have a positive and significant impact on that of large firms, and large firms in a specific industry do not show any tendency to imitate small firms' disclosure practices. The coefficient β2 in the regression model captures the peer-imitation effect in ESG disclosure between large and small firms within the same industry. Given its non-significance, the fourth hypothesis is not rejected.

4.1.4. Testing of the H5, H6

To test the fifth and sixth hypotheses, Regression Model (3) is estimated for the sample of small firms. The results from estimating Regression Model (3) for testing these hypotheses are presented in Table 5.

Table 5. Regression results for H5, H6

$ESG_{i,t} = \alpha + \beta_1 B\ ESG_{Industry,t-1} + \beta_2 S\ ESG_{Industry,t-1} + \beta_3 CR_{i,t} + \beta_4 LEV_{i,t} + \beta_5 ROA_{i,t} + B_6 QTobin_{i,t} + \beta_7 Grow_{i,t} + \beta_8 Auditor\_Size_{i,t} + \beta_9 Opin_{i,t} + \varepsilon_i$

Variable	Coefficient	Std. Error	T-Statistic	P-Value
B $ESG_{Industry}$	0.396	0.147	2.682	0.007
S $ESG_{Industry}$	0.303	0.168	1.802	0.072
CR	2.716	1.830	1.483	0.138
LEVRAGE	0.696	0.656	1.061	0.289
ROA	0.996	2.122	0.469	0.639
Q TOBIN	-0.109	0.047	-2.331	0,020
GROW	-0.003	0.021	-0.181	0.856
$Auditor_{Size}$	-0.891	0.548	-1.623	0.105
OPIN	-1.474	0.961	-1.534	0.125
C	3.823	2.471	1.546	0.122
Year & Industry Effect	yes			
Adjusted R^2	0.166			
F-Statistic (Prob)	9.619 (0.000)			

According to the fifth hypothesis, in small firms, there is an imitative intra-industry peer effect on ESG disclosure relative to large firms. The significance level for the ESG disclosure score of large firms in the industry (B $ESG_{Industry}$) is 0.0076, which is below the 5% error threshold. Additionally, the coefficient for this variable is 0.3976, indicating that ESG disclosure by large firms in the industry has a positive and significant impact on the ESG disclosure of small firms. In other words, an imitative intra-industry peer effect exists in ESG disclosure between small and large firms within the same industry, meaning that small firms engage in informational imitation of large firms. Therefore, the fifth hypothesis is not rejected at the 95% confidence level, suggesting that ESG disclosure by large firms in an industry enhances ESG disclosure by small firms.

According to the sixth hypothesis, in small firms, there is a reciprocal intra-industry peer effect on ESG disclosure. This hypothesis expected the coefficient of the variable $S_{ESGIndustry}$ to be positive and significant. However, based on the results in Table 5, the coefficient's significance level is 0.0722, which is greater than the 5% error threshold. Therefore, the sixth hypothesis is rejected at the 95% confidence level. In other words, ESG disclosure by small firms in a given industry does not have a positive and significant effect on the ESG disclosure of other small firms. This implies that small firms are not inclined to engage in informational imitation of their small-firm peers within the same industry.

5. Discussion and conclusions

Given the significance of peer effects across firms and industries, this study examines the intra-industry peer effect on ESG information disclosure, with a specific focus on the moderating roles of firm size and experience. By doing so, the study aims to provide deeper insights into the mechanisms driving corporate reporting behaviors in response to peer influence.

The results show that there is a large set of intra-industry peer effects in ESG disclosure level between firms, which is in line with the literature review done by the study. Alternatively, companies follow their industry peers' lead in ESG disclosure. which indicates that industry-level ESG disclosure has a positive and statistically significant impact on firm-level ESG disclosure. Firms' strategic imitation of industry leaders to remain relevant and meet their social responsibility, in turn, leads to a peer effect in this regard. These findings are consistent with previous studies by Hu et al. (2022) and Joodaki et al. (2023).

Drawing on this, the findings show that firm experience strengthens the peer effect on ESG disclosure, contrary to previous studies' findings. Moreover, because firms follow their peers' ESG disclosure behavior, as they gain more experience, they continue to do so. This behavior aligns with information-based imitation theory (Lieberman and Asaba, 2006), which posits that firms, regardless of their experience level, imitate others to address asymmetric information. Mimetic behavior is typical among social creatures; thus, it should come as no surprise that both functionalist and interactionist sociologists also theorize it across most social groups, including corporate entities. However, these results are at odds with Elahi et al. (2020), Wan et al. (2016), Zhi-guo and Bo-yi (2020), and Chen and Ma (2017).

The results also reveal an intra-industry peer effect in ESG disclosure, especially among large firms, which aligns with the theoretical constructs underpinning the current study. Larger companies replicate the ESG disclosure practices of similarly large peers in their industry. The desire for industry leadership, stronger stakeholder trust, better borrowing terms, and improved sustainability presentation drives such imitative behavior. These findings are consistent with the notion of reciprocal peer effects, which refers to the impact one firm has on a similar-sized firm (Hu et al., 2022), and they echo previous research by Hu et al. (2022), Roberts and Leary (2014), and Haunschild and Miner (1997).

Furthermore, the findings indicate that large firms do not exhibit an imitative peer effect toward ESG disclosures made by small firms, lending further credence to the paper's hypothesis. This indicates that big corporations are less willing to emulate ESG practices from industry peers on a smaller scale. Indeed, such asymmetric imitation behavior aligns with the accounting literature, which shows that for-profit and not-for-profit peer firms exhibit a greater propensity to imitate those they believe are better or more visible, doing so in ways necessary to respond positively to stakeholder expectations (or to emulate peers with good qualities). Hu et al. (2014) and Leary et al. (2022) further support these findings.

Likewise, the results reveal a peer effect on ESG disclosure within industries, with small firms mimicking large firms. This finding aligns with the existing literature. It supports the imitative peer effect and information-based imitation theories, in that small firms are more likely to copy industry leaders' policies to gain access to better information and build their credibility. These results align with previous studies conducted by Hu et al. (2022) and Chen and Ma (2017).

Finally, the evidence suggests that scaling up peer-group influence amongst smaller firms does not appear to result in reciprocal behavior on ESG disclosures, as expected, but rather in mimicking larger-firm behaviors (Hu et al., 2022), which differs from Leary and Roberts (2014). In this section, several theoretical and contextual factors are presented to frame the overall findings that explain this absence. According to information-based theory, small firms tend to regard larger firms (which possess superior resources and knowledge) as central actors in information networks and, as a result, induce imitative peer effects (Lieberman and Asaba, 2006; Li et al., 2018). This aligns with social learning theory, small firms are more likely to imitate leaders than peers (Bikhchandani et al., 1998). In line with this, competition-based theory suggests that small firms mimic large ones to conform to sector expectations and that they have weak competitive drivers among themselves in the ESG fields, which mitigates reciprocal impacts (Wan et al., 2016; Cho et al., 2025). Due to resource scarcity (Haunschild and Miner, 1997), small firms typically prioritize operational needs that are less mutually aligned between the acting firm and the imitator, thereby diminishing reciprocal imitation. In addition, heterogeneity in business models and environmental, social & governance (ESG) priorities among small firms renders peer disclosures less valuable to them; they are more likely to mimic the practices of large firms. All of these factors combine to demonstrate the interplay between theoretical and contextual dynamics, which will also explain why reciprocal peer effects do not persist in Iran's small firms.

The implications of these findings are highly relevant for policymakers interested in improving corporate ESG reporting. Governments should use cross-industry imitation to lift the average quality of ESG disclosures. Specifically, policymakers should clearly identify and promote firms with exemplary ESG disclosure practices as an industry benchmark. These model firms should be publicly recognized through media campaigns and incentivized with long-term tax benefits. Establishing industry leaders can provide a framework for firms with weaker ESG practices, guiding them on what to disclose and how. Regulatory bodies should also consider providing standardized frameworks for integrated ESG reporting to ensure consistency and comparability. Additionally, enforcing compliance with ESG reporting standards can foster active learning and imitation among firms. It is crucial to account for firm size and experience when promoting ESG reporting, as these factors significantly influence intra-industry imitation.

For future studies, it is recommended to examine the impact of peer financial reporting quality on peers' ESG disclosure. Additionally, by analyzing the peer effects of ESG disclosure across industries, it can be investigated which industries experience stronger peer effects.

Reference

1. Adhikari, B. K. and Agrawal, A. (2018). Peer influence on payout policies. *Journal of Corporate Finance*, 48, pp. 615-637. https://doi.org/10.1016/j.jcorpfin.2017.12.010
2. Akbas, H. E. (2014). Company characteristics and environmental disclosure: an empirical investigation on companies listed on Borsa Istanbul 100 index. *The Journal of Accounting and Finance*, 62, pp. 145-164. https://dergipark.org.tr/tr/download/article-file/427507
3. Bahramoghadam, M., Sadeghi, Z.A. and Safarzadeh, S. (2013). Investigating the relationship

between corporate governance mechanisms and corporate social responsibility disclosure, quarterly. *Journal of Financial Accounting*, 5(20), pp. 90-107 (In Persian). http://qfaj.mobarakeh.iau.ir/article-1-123-fa.html

4. Bernardi, C. and Stark, A. W. (2018). Environmental, social and governance disclosure, integrated reporting, and the accuracy of analyst forecasts. *The British Accounting Review*, 50(1), pp. 16-31. https://doi.org/10.1016/j.bar.2016.10.001
5. Bikhchandani, S., Hirshleifer, D. and Welch, I. (1998). Learning from the behavior of others: Conformity, fads, and informational cascades. *Journal of Economic Perspectives,* 12(3), pp. 151-170. https://doi.org/10.1257/jep.12.3.151
6. Bonaiuto, M., Fornara, F. and Bonnes, M. (2003). Indexes of perceived residential environment quality and neighborhood attachment in urban environments: a confirmation study on the city of Rome. *Landscape and Urban Planning*, 65(1-2), pp. 41-52. https://doi.org/10.1016/S0169-2046(02)00236-0
7. Capelle-Blancard, G. and Petit, A. (2019). Every little help? ESG news and stock market reaction. *Journal of Business Ethics*, 157(2), pp. 543-565. https://doi.org/10.1007/s10551-017-3667-3
8. Chen, S. and Ma, H. (2017). Peer effects in decision-making: Evidence from corporate investment. *China Journal of Accounting Research*, 10(2), pp. 167-188. https://doi.org/10.1016/j.cjar.2016.11.002
9. Cho, H., Choi, A. and Jung, T. (2025). Peer effects in investment: Evidence from early-tenure CEOs. *European Accounting Review,* 34(3), pp. 1249-1275. https://doi.org/10.1080/09638180.2024.2363444
10. Demerjian, P., Lev, B. and McVay, S. (2012). Quantifying managerial ability: A new measure and validity tests. *Management Science*, 58(7), pp. 1229-1248. https://doi.org/10.1287/mnsc.1110.1487
11. DiMaggio, P. J. and Powell, W. W. (2022). The iron cage revisited: collective rationality and institutional isomorphism in organizational fields. *American Sociological Review,* 48(2), pp. 145-160. http://www.jstor.org/stable/2095101
12. Dodgson, M. (1993). Organizational learning: a review of some literature. *Organization studies*, 14(3), pp. 375-394. https://doi.org/10.1177/017084069301400303
13. Duong, H. K., Ngo, A. D. and McGowan, C. B. (2015). Industry peer effect and the maturity structure of corporate debt. *Managerial Finance*, 41(7), pp. 714-733. https://doi.org/10.1108/MF-02-2014-0050
14. Eccles, R. G. and Serafeim, G. (2013). The performance frontier. *Harvard Business Review*, 91(5), pp. 50-60.
15. Elahi, E., Khalid, Z. and Zhang, Z. (2022). Understanding farmers' intention and willingness to install renewable energy technology: A solution to reduce the environmental emissions of agriculture. *Applied Energy*, 309, Article 118459. https://doi.org/10.1016/j.apenergy.2021.118459
16. Fakhari, H., Malekian, E. and Jafaei-Rahni, M. (2017). Explanation and ranking of environmental, social and corporate governance components using the Analytic Hierarchy Process in companies listed on the Stock Exchange. *Quarterly Journal of Value and Behavioral Accounting*, 2(4), pp. 153-187 (In Persian). https://doi.org/10.29252/aapc.2.4.153
17. Gimeno, J., Hoskisson, R. E., Beal, B. D. and Wan, W. P. (2005). Explaining the clustering of international expansion moves: A critical test in the US telecommunications industry. *Academy of Management Journal*, 48(2), pp. 297-319. https://doi.org/10.5465/amj.2005.16928411
18. Golijani, M. and Jafar Salehi, K. (2024). Investigating the relationship between environmental

disclosure, earnings management and accounting conservatism. *Contemporary Research in Management and Accounting Sciences,* 16(5), pp. 77-93 (In Persian). https://civilica.com/doc/1669632
19. Grewal, J., Hauptmann, C. and Serafeim, G. (2021). Material sustainability information and stock price informativeness. *Journal of Business Ethics*, 171(3), pp. 513-544. https://doi.org/10.1007/s10551-020-04451-2
20. Gu, C. and Shen, M. (2024). Peer effect of environmental information disclosure behavior in China: Empirical test based on" Real green" and" Fake green". *Heliyon*, 10(23), e40779. https://doi.org/10.1016/j.heliyon.2024.e40779
21. 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
22. Han, J. C., Wild, J. J. and Ramesh, K. (1989). Managers' earnings forecasts and intra-industry information transfers. *Journal of Accounting and Economics*, 11(1), pp. 3-33. https://doi.org/10.1016/0165-4101(89)90012-8
23. Haunschild, P. R. and Miner, A. S. (1997). Modes of interorganizational imitation: The effects of outcome salience and uncertainty. *Administrative Science Quarterly*, 42(2), pp. 472–500. https://doi.org/10.2307/2393735
24. Hu, K., Hao, Y. and Yu, D. (2022). Intra-industry peer effect in corporate environmental information disclosure: Evidence from China. *Plos One*, 17(9), pp. 1-19. https://doi.org/10.1371/journal.pone.0274787
25. Joodaki, M., Khalili, Y. and Rahmati, L. (2023). Investigating the moderating effect of Leader and Young firms in industry on the relationship between Peer Firms and the Firm's Investment based on the theories of Information Deficit and Competitiveness. *Journal of Accounting Advances,* 15(2), pp. 93-124. https://doi.org/10.22099/JAA.2024.48344.2369
26. Kaustia, M. and Rantala, V. (2015). Social learning and corporate peer effects. *Journal of Financial Economics*, 117(3), pp. 653-669. https://doi.org/10.1016/j.jfineco.2015.06.006
27. Khozein, A., Talebnia, G., Garkaz, M. and Banimahd, B. (2018). Effect of ownership structure on the development of the sustainable reporting. *Management Accounting Research Quarterly*, 11(36), pp. 69-101 (In Persian). https://sanad.iau.ir/Journal/jma/Article/816328
28. Kordestani G. and Jafari Souq S. A. (2022). Peer firm performance and earnings management: the effect of capital market pressure. *The Journal of Judgment, and Decision Making in Accounting and Auditing*, 1(3), pp. 71-9 (In Persian). 10.30495/jdaa.2022.693174
29. Leary, M. T. and Roberts, M. R. (2014). Do peer firms affect corporate financial policy? *The Journal of Finance,* 69(1), pp. 139-178. https://doi.org/10.1111/jofi.12094
30. Li, Z. S., Su, C. and Li, H., Dongmin, K. (2018). Excess leverage and region-based corporate peer effects. *Journal of Financial Research*, 1(9), pp. 74-90.
31. Lieberman, M. B. and Asaba, S. (2006). Why do firms imitate each other? *Academy of Management Review*, 31(2), pp. 366-385. https://doi.org/10.5465/amr.2006.20208686
32. Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, 60(3), pp. 531-542. https://doi.org/10.2307/2298123
33. Wan, L., Liang, C. and Rao, J. (2016). Industry peer effect in M&A decisions of China's listed companies. *Nankai Business Review*, 19(3), pp. 40-50.
34. Wang, Y. T., Wang, S. and Duan, M. Y. (2024). The peer effect of green investment activities of China's listed firms. *Systems Engineering Theory and Practice*, 44(3), pp. 874-892. https://doi.org/10.12011/SETP2023-0011
35. Xi, L., Bian, H. and Wang, X. (2024). Is there a peer effect on corporate ESG performance?

Evidence from China's capital market. *Chinese Journal of Population, Resources and Environment*, 22(4), pp. 423-435. https://doi.org/10.1016/j.cjpre.2024.11.006

36. Yang, S., Ye, H. and Zhu, Q. (2017). Do peer firms affect firm's corporate social responsibility? *Sustainability*, 9(11), p. 1967. https://doi.org/10.3390/su9111967
37. Zhang, Z. and You, J. (2024). Does firms' ESG information disclosure have contagion effect? Evidence from China. *Corporate Social Responsibility and Environmental Management*, 31(4), pp. 3274-3296. https://doi.org/10.1002/csr.2737
38. Zhi-guo, D. I. N. G. and Bo-yi, L. I. (2020). Regional peer effect of dividend policy of listed companies. *Journal of South China Normal University (Social Science Edition)*, 1(3), pp. 95-107.

IJAAF; Vol. 10, No. 2, Spring 2026, Serial 35

Ferdowsi University of Mashhad

Iranian Journal of
Accounting, Auditing & Finance

Quarterly

RESEARCH ARTICLE

The Impact of Conspiracy Illusion on Auditor Independence: The Moderating Roles of Tolerance for Ambiguity and Religious Orientation

Shirin Molavi, Ameneh Bazrafshan *

Department of Accounting, Imam Reza International University, Mashhad, Iran

How to cite this article:
Molavi, S. and Bazrafshan, A. (2026). The Impact of Conspiracy Illusion on Auditor Independence: The Moderating Roles of Tolerance for Ambiguity and Religious Orientation. Iranian Journal of Accounting, Auditing and Finance, 10(2), 55-67. doi: 10.22067/ijaaf.2026.47627.1585
https://ijaaf.um.ac.ir/article_47627.html

ARTICLE INFO

Article History
Received: 2025-05-03
Accepted: 2025-08-15
Published online: 2026-04-12

Abstract

Auditor independence is essential to reliable financial reporting, yet prior research has largely focused on structural, regulatory, and economic determinants, with limited attention to the psychological factors that may influence auditors' professional judgment. To address the gap, this study investigated how conspiracy illusion (the socio-psychological predisposition to see hidden motives and manipulations) influences auditor independence. Data from auditors in Iran, a newly emerging country molded by socio-cultural and religious factors that can impact ethical decisions, is used to expand upon the moderating effects of tolerance for ambiguity and religious orientation. Additionally, examining psychological factors in large countries like Iran, where corporate governance is weak and a more permissive natural environment allows for greater individual expression, offers valuable insights. Using data from 300 auditors analyzed through structural equation modeling (SEM), the findings revealed a moderate but significant negative relationship between conspiracy illusion and auditor independence—higher levels of conspiracy illusion are associated with lower independence. However, auditors with greater tolerance for ambiguity and stronger religious orientation experienced a weakened negative effect of conspiracy illusion. These findings advanced understanding of auditor independence by demonstrating its shaping by professional and regulatory standards and underlying psychological and cultural factors. Consequently, fostering psychological resilience, critical thinking, and positive ethical and religious attitudes may help strengthen auditor independence, particularly in complex socio-cultural contexts like Iran.

Keywords:
Auditor Independence, Conspiracy Illusion, Religious Orientation, Tolerance for Ambiguity

https://doi.org/10.22067/ijaaf.2026.47627.1585

CC BY

NUMBER OF REFERENCES
27

NUMBER OF FIGURES
2

NUMBER OF TABLES
3

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

*Corresponding Author: Ameneh Bazrafshan
Email: bazrafshan_63@yahoo.com
Tel: 09151337566
ORCID:

1. Introduction

Today's business and professional environment is increasingly complex, making data-driven governance and decision-making a top priority. Out of all types, financial data is the most essential to ensure organizational accountability and bolster stakeholder trust. However, there is an inherent conflict of interest between those preparing financial data (suppliers) and those who rely on it (users), giving reasons to be concerned about the credibility of such data. Auditors help fill this gap by attesting to the appropriateness and reliability of financial statements (Ghaznavi Doozandeh et al., 2022). Despite extensive research on regulatory and economic determinants of auditor performance, the influence of auditors' psychological factors on their independence remains underexplored, particularly in emerging markets. This study addresses this gap by examining how conspiracy illusion, along with tolerance for ambiguity and religious orientation, affects auditor independence, thereby extending prior literature on the psychological and socio-cultural determinants of audit quality.

Nevertheless, a fundamental prerequisite for auditors to perform their roles is effective independence. Auditor independence is an abstract concept which cannot be observed directly and is generally construed as a mental state by which objectivity, integrity, trustworthiness, and evaluative judgment seem to exist. Independence of the auditor is necessary because independence cannot be compromised due to financial, political or personal relationships with the client or employer (Citron 2003). Because of its key importance, researchers have tried for a long time have tried to find what shapes the independence of auditors. Previous research has focused extensively on regulatory and economic drivers of decay, while psychological factors that contribute to decay are more poorly understood. This research filled that need by looking at conspiracy illusion as a new social and psychological phenomenon that could compromise auditors' objectivity on the job. Understanding the effect of conspiracist ideation is particularly essential, since auditors' cognitive biases and susceptibility to unfounded beliefs may necessarily distort their objective judgment insidiously, damaging financial reporting quality and stakeholders' confidence.

Conspiracy illusion is defined as the belief that major events or social problems are orchestrated behind the scenes by powerful people or organizations (Swami et al., 2010). This worldview leads to the inclination to accept weak or unverifiable evidence and undermines skepticism in regard to veritable sources of information. People may be led to blame others instead of recognizing their own role or working towards correcting the course (Biddlestone et al., 2020). Conspiracist ideation is no trifling matter; it represents serious and consequential potential (Douglas et al., 2019). Examples include hypersensitivity, distrust of others, defensiveness, overreacting to what can merely appear as criticism, taking neutral remarks personally, being hostile and argumentative or rigid/inflexible and domineering, thinking that others are not capable of forgiveness; paranoid about the false information spread by others behind one's back, believing that other people cannot be honest in their dealings with you and many times counter accusing of everyone who meet (Swami et al., 2010; Douglas and Sutton, 2018). An understanding of these characteristics is especially meaningful in the realm of auditing, as such cognitive biases can lead to impairment of professional judgment and affect objectivity.

Moreover, distrust and cynicism (Swami et al., 2010), Machiavellianism (Douglas and Sutton, 2018), and collective narcissism (De Zavala et al., 2009) were found to be positively related to conspiracist ideation (Douglas et al., 2019). This need is particularly common in situations associated with feelings of powerlessness (Whitson and Galinsky, 2008) and uncertainty (Van Prooijen and Jostmann, 2013). Although conspiracist ideation is typically observed in a negative light, some positive suggestions have also been extracted from previous research, including bridging social awareness, critiquing power dynamics or promoting innovation and creativity. Diverse views on its effects on auditor independence can be related to the established theories from psychology. Under the locus of control perspective, people with high conspiracist ideation are likely to link their financial success or failure to external entities as opposed to their own behaviors and actions, which reduces self-efficacy and motivation and ultimately impairs auditor independence. On the other hand, professional skepticism theory proposes that a general disbelief in others and the notion that certain individuals or groups of people may not reach an expected standard, causing auditors to be more independent within their own evaluations. Against this backdrop of disparate theoretical predictions, the current study empirically investigates how individual-level conspiracist ideation relates to auditor independence. Moreover, the research examines two socio-psychological moderators — tolerance for ambiguity and religious orientation. Acceptance of ambiguity, as noticed in the article title, refers to humans' ability to endure ambiguous situations and accept uncertainty without experiencing extreme anxiety or discomfort. Auditors are often faced with inconsistent evidence, missing information, and multiple plausible outcomes, thus highlighting the importance of ambiguity tolerance for judgment quality (e.g., conservatism) and decision-making under uncertainty. Religious orientation might also affect auditors' decision-making and behaviors, and will positively affect professional skepticism of auditors or perception of risk related to material misstatements. According to existing research, audit performance can be influenced by auditors' worldwide motivation and based on their religious alignment, which, in turn, makes them perform better at work. From multiple angles, this study helps expand knowledge and, indeed, the existing literature. First, it is the first study to explore the effect of conspiracist ideation on auditor independence, thus providing empirical evidence for the theory suggesting a negative relationship between conspiracist ideation and independence. Second, the results showed that promoting tolerance for ambiguity and a positive religious orientation is possible, which can reduce the negative effect of conspiracist ideation on auditor independence.

2. Theoretical background and literature review

Independence is considered a key aspect of the auditing profession to guarantee credibility, objectiveness and unbiasedness in audit reports (Data until October 2023). In addition, independence is the auditor's ability to execute their professional duties without influence by personal interests, client management decisions, or other affiliations that can interfere with judgment. In literature, independence is separated into independence, including the auditor's ability to make impartial judgment and independence in appearance, the perception that each stakeholder who is interested in auditors is objective (Liu et al., 2014). Independence is vital for public confidence and the smooth functioning of financial markets, with poor independence associated with high-profile corporate failures, undetected fraud and mismanagement. Theoretical views contribute to the understanding of auditor independence determinants. According to agency theory, auditors are needed as monitors within companies to help diminish conflicts that may arise between the principals (usually referred to as shareholders) and agents (management). Thus, emphasizing the auditors' importance of professional objectivity and ethical vigilance. From a moral psychology standpoint, auditors’ ethical reasoning, integrity and susceptibility to cognitive biases drive their ability to be independent under pressure. Along similar lines, research on cognitive biases has found that psychological tendencies like overconfidence, confirmation bias and susceptibility to conspiratorial thinking can exert a subtle influence over auditors’ judgments and decisions. These theoretical lenses highlight the importance of examining individual-level psychological factors, in addition to structural and regulatory mechanisms, to explain differences in auditor independence. In this context, a novel and plausible contributor has been identified as conspiracist ideation or the proclivity to assume that significant events in the world are being secretly orchestrated by powerful others (Swami et al., 2010).

Conspiracist ideation might affect auditor behavior by coloring information processing, attribution of responsibility, and trust in clients' financial reporting, which are integral elements of professional judgment. Auditors who fall prey to conspiracist logic may thus either become doubters (so extremely that they see all evidence as potentially fake) or externalize responsibility (attributing failures not to missteps of their own, but rather "pulling the strings" behind-the-scenes actors). Given theories of locus of control and professional skepticism, this dual potential influence, whether positive or negative, has complex effects upon auditor independence. Prior empirical investigations have examined related constructs like distrust (Swami et al., 2010), cynicism (Douglas and Sutton, 2018), Machiavellianism (De Zavala et al., 2009) and collective narcissism. These results suggest that personality traits and cognitive styles may heavily influence auditors' evaluations and risk judgments. However, to the best of our knowledge, no studies have systematically investigated how conspiracist ideation interacts with moderators, such as tolerance for ambiguity and religious orientation, to promote auditor independence. This research gap is critical to understand and find ways to improve objectivity in audit decision-making through an examination of the socio-psychological determinants of audit quality.

The term 'conspiracy' ideation describes a more trait-like orientation, defined as the propensity to explain major events or social phenomena with secret plots by powerful collectives acting out of self-interest (Douglas et al., 2019). Conspiracy ideation is therefore a general cognitive and motivational style marked by a distrust of those around us, an external perception of control, and ease towards simple explanations for complex situations (Douglas, 2021) rather than mere aggregation of idiosyncratic beliefs. Such an orientation often emerges in attempts to regain order or predictability in conditions of uncertainty, threat or lack of control (Whitson and Galinsky, 2008). At the individual level, conspiracist ideation shares territory with paranoia and mistrust, which includes exaggerated perceptions of manipulation, persecution, or hidden intent. These inclinations can vary from mild to severe paranoia and are often associated with dimensions of interpersonal sensitivity, perceived persecution and the external control beliefs (Freeman et al. 2011). Sustained exposure to this line of reasoning correlates with greater social distrust, cognitive rigidity and interpersonal dysfunction. In behavioral terms, those who engage with conspiratorial thinking tend to ascribe events to sinister undercurrents (Sunstein and Vermeule, 2009)—malevolently motivated agencies or institutions that conspire behind the scenes. This way of thinking provides cognitive coherence to individuals and helps them make sense out of ambiguous or threatening information (Douglas, 2021) and breeds hostility, cynicism, and authoritarianism (Swami et al., 2010). The consequences of conspiracist ideation are theoretically ambivalent in the auditing context, which can increase auditor independence by taking vigilance, skepticism and the sensitivity to possible manipulation one step further on the other. Such auditors might challenge evidence more rigorously and process information at greater depth, thereby lowering scrutineering risk. However, being overly indulgent in a conspiratorial mindset can threaten independence by strengthening confirmation bias and social distrust while reading too much into ambiguous signals paralyzing objectivity and professional communication (Gunawan and Lestari, 2025). Such an attitude, when sustained, also yields risks of being more strenuous psychologically, reduced intra-team cooperation, as well as a risk of losing client trust, all three negatively influencing the quality of the audit (Elinda et al., 2019). Such mixed outcomes have been backed up with empirical studies. According to Khaleq (2023), psychological strain and cognitive bias applied to auditors affect their independence and ethical consistency. Likewise, Mardijuwono and Subiyanto (2018) found that professional skepticism responding with negative ideological dispositions, such as conspiracist ideation, yields imbalanced perspectives and alienated assessments.

Based on these opposing theories, the first hypothesis of the study is formulated as follows:

Building on the preceding discussion, so that conspiracist ideation may meaningfully shape auditors' professional behavior and judgment. People who exhibit higher levels of conspiratorial thinking are also more skeptical and vigilant, traits that can bolster or undermine independence depending on how they are expressed. Skepticism, when used with a critical evaluation process, may help in making an objective, less biased assessment. However, if it is based on distrust and cognitive rigidity, it can reinforce biased reasoning, selective information processing and impaired collaboration, all of which put independent judgment at risk. Accordingly, understanding the direction and magnitude of this influence is critical for clarifying the psychological underpinnings of auditor independence.

H1: Conspiracy illusion has a significant relationship with auditor independence.

In theory, tolerance for ambiguity can moderate the relationship between conspiracist ideation and auditor independence. This construct describes an individual's ability to cope with partial, inconsistent or ambiguous information as a key factor in decision-making and professional judgment. This phenomenon, initially coined by Frenkel-Brunswik (1949) in a research (albeit also related to uncertainty), has grown over the past several decades into an immune psychological dimension of how humans interact with uncertainty. Instead of worrying about the word's etymology, one should consider tolerance for ambiguity as a cognitive-emotional capacity that allows one to process conflicting evidence, without excessive anxiety or premature closure. This trait is crucial in the auditing context since auditors frequently work in environments with limited amounts of data, time pressure, and ambiguity with respect to the audit evidence. A high tolerance of ambiguity enables deliberate thought, balanced skepticism and resistance to cognitive biases that can undermine audit judgments. On the other hand, low tolerance for ambiguity could create rigidity and anxiety or encourage simplistic or conspiratorial approaches to complex audit issues. High tolerance for ambiguity predicts lower anxiety in the face of incomplete or contradictory information and superior performance in changing situations. For auditors, this means greater ability to integrate disparate pieces of evidence, exercise professional skepticism, and maintain independence under uncertain or politically charged audit conditions. In contrast, people low in need for closure may want certainty and come to adopt whatever beliefs confirm their assumptions because they are sure not to be disappointed. When these tendencies are conjoined with conspiracist ideation, they can exacerbate distrust and lower receptiveness to alternative explanations that would otherwise enhance impartial judgment. As such, tolerance for ambiguity is anticipated to moderate the association between conspiracist ideation and auditor independence. High ambiguity tolerance auditors would be able to preserve the independence of thought in the face of conspiracy cues, while low ambiguity tolerance would succumb to biased evidential reasoning and excessive skepticism that erodes. Empirical evidence supports this theoretical reasoning. The positive effect of the biases of ambiguity aversion and responsibility dimension on audit quality was confirmed in a study conducted by Karimi et al. (2021). However, overconfidence bias and optimism and extraversion variables negatively affect audit quality.

High tolerance of ambiguity allows auditors to effectively minimize their exposure to affected CPAs, remaining independent and adhering to skeptical evaluation techniques. On the other hand, low tolerance of ambiguity is related to discomfort in ambiguous contexts, a cognitive style relying on fixed categories and rigidity, and increased susceptibility to conspiracist ideation. Therefore, it is expected that tolerance for ambiguity moderates the relationship between conspiracist ideation and auditor independence—in particular, higher levels of tolerance are associated with greater independence (vs. lower), while lower levels of tolerance are associated with reduced independence. Accordingly, the second research hypothesis states:

The theoretical discussion above suggests that increased tolerance for ambiguity provides a buffer against the negative effects of conspiracist ideation on auditor independence by inducing cognitive flexibility and reducing bias in conditions of uncertainty.

H2: The level of tolerance for ambiguity has a significant effect on the relationship between conspiracy illusion and auditor independence.

Religious psychology studies the psychological effects of religion, as it has always played a major role in shaping human cognition and behavior. The literature has shown that although following any value system contributes to mental integration and psychological coherence, religious doctrines can offer a platform for moral consistency and psychological balance because they touch upon existential subjects (such as suffering, death, and meaning) and ethical standards (such as kindness, wisdom and self-denial) (Abu Raiya et al., 2008).

Religious attitudes can also help individuals view the world with hope and trust in others. Those with strong religious principles may be less prone to conspiracy illusion and instead emphasize moral principles and honesty.

Religious teachings typically stress moral values such as honesty, fairness, and trust. Individuals with such attitudes may, when faced with conspiracy illusion, be inclined to focus on moral principles and turn toward proper and fair behavior rather than relying on conspiracy theories. Auditors with sound and positive religious attitudes are usually committed to moral principles. This may help them maintain their independence and make precise, unbiased judgments when faced with pressures arising from conspiracy illusion.

In addition, religious attitudes can reinforce an auditor's confidence in their capabilities and soundness of judgment. Their confidence in their actions frees them from the influence of conspiracy theories. In summary, religious attitudes are a moderating variable in the relationship between conspiracy illusion and auditor independence. Individuals with positive, law-abiding religious attitudes are more likely to preserve their independence in the face of conspiracy illusion and act according to ethical and professional principles. In contrast, individuals with conspiracy illusion who lack religious principles may be influenced by negative thinking and lose their independence.

Duh et al. (2022) found that the religious illusion of audit partners has a positive and significant impact on audit quality. Ezzati Shalkoni et al. (2024) found that the personal characteristics of auditors have a positive and significant impact on religious attitudes and audit quality. Azhdary et al. (2025) concluded that the variables of ethical principles and religiosity of auditors have a significant effect on the quality of services provided by independent auditors and lead to an increase in audit quality. Therefore, religious beliefs are expected to lessen the negative effects of conspiracy theories on an auditor's independence because these beliefs strengthen moral self-control, ethical thinking, and the ability to resist cognitive biases.

H3: Religious orientation significantly moderates the relationship between conspiracy illusion and auditor independence.

3. Methodology

This study used a descriptive–correlational research design and a survey-based approach to collect quantitative data. The research was framed against a deductive–inductive reasoning approach, where theories were formulated through their theoretical underpinnings and subsequently tested in practice. The statistical population consists of all auditors who worked in the Iranian Audit Organization and member firms of the Iranian Association of Certified Public Accountants (IACPA). The population included auditors of different professional levels (trainee, auditor, senior auditor, audit manager) who were involved in the auditing profession. To ensure the sample was representative, auditors were stratified based on their professional rank and where they worked, using a stratified random sampling method. In 2025, a total of 450 questionnaires were distributed. After removing invalid responses, 300 valid ones were left for analysis, resulting in a response rate of about 67%. Stratification allowed for a proportional distribution across gender, job level and membership status in IACPA, which improved generalizability. Data collection was performed using a five-part questionnaire:

1. Demographic data: Auditor's rank, gender and professional experience as auditing – Institute of accountants and auditors professionals' age level.
2. Conspiracist Ideation: nine items, 5-point Likert scale (adapted from Jolley et al. (2021).
3. Ambiguity Tolerance: 13 items with a 5-point Likert scale, adapted from McLain's (2009) canonical questionnaire.
4. Religious Orientation: 25 items based on Golriz's (1974) scale
5. Auditor Independence: Nine out of 14 items assessed Kurnia's (2021) validated scale.

To test the structural model, demographic and occupational characteristics (gender, professional experience, job rank) were controlled for potential confounding effects. Lastly, Partial Least Squares Structural Equation Modeling (PLS-SEM) is used for hypothesis testing in SmartPLS 3 software, due to its robustness for complex models and small sample sizes.

4. Results and data analysis

Demographic characteristics of the study participants are shown in Table 1. The majority of the sample were male (74.3%), while females accounted for 25.7%. The majority of respondents were under 35 years old and about one-third had a master's degree. Most were employed in auditor or senior auditor positions, and a minority were managers or supervisors. Notably, 18.7% of participants were certified auditors with varying degrees of professional qualification.

To improve the interpretability of forthcoming analyses, primary demographic variables (gender, age, education, and certification status) were tested to determine if their effects were substantive and statistically controlled for in the main and moderation analyses. This method allows for the analysis of links between conspiracist ideation, tolerance for ambiguity, religious orientation and auditor independence without those relationships being confounded by demographic influences.

Table 1. Demographic characteristics of the survey participants

Demographic Characteristic	Category	Frequency	Percentage
Gender	Female	77	25.7%
	Male	223	74.3%
Age Group	20–25 years	78	26.0%
	25–30 years	74	24.7%
	30–35 years	48	16.0%
	35–40 years	51	17.0%
	Above 40 years	49	16.3%
Educational Level	PhD	21	7.0%
	Master's degree	121	40.3%
	Bachelor's degree and below	158	52.6%
Job Position	Auditor	164	54.6%
	Senior Auditor	39	13.2%
	Audit Supervisor	32	10.6%
	Audit Manager	40	13.3%
	Partner in Firm	25	8.3%
Membership in the Official Auditors' Society	Yes	56	18.7%
	No	244	81.3%
Total		300	100%

Model Fit Indices: To assess the model's validity and fit, several statistical indicators were used:

The coefficient of determination (R^2), standardized path coefficients (T-values), composite reliability (CR), average variance extracted (AVE), standardized root mean square residual (SRMR), and Chi-square statistics. The overall model goodness-of-fit (GOF) was 0.658, indicating a strong fit based on threshold benchmarks (weak = 0.01, moderate = 0.25, strong = 0.36).

Other fit indicators also support the model's adequacy:

- SRMR = 0.028
- $R^2 = 0.683$
- Normed Fit Index (NFI) = 0.912
- Chi-Square = 462.860

In assessing the construct validity, both convergent validity and discriminant validity were assessed. All constructs satisfied convergent validity (AVE > 0.5) and composite reliability (CR > 0.7), as summarized in Table 2. To assess for discriminant validity, the square root of the AVE of each construct was greater than its correlations with other constructs, supporting that variables are indeed different from one another (as shown in Table 3).

Table 2. Convergent validity and reliability of constructs

Construct	Cronbach's Alpha	Composite Reliability (CR)	Average Variance Extracted (AVE)
Auditor Independence	0.938	0.974	0.644
Conspiracy Illusion	0.904	0.922	0.570
Tolerance for Ambiguity	0.905	0.965	0.549
Religious Orientation	0.962	0.982	0.529

Table 3. Discriminant validity of constructs (Fornell–Larcker Criterion)

Variable	Conspiracy Illusion	Auditor Independence	Tolerance for Ambiguity	Religious Orientation
Auditor Independence	0.803	—	—	—
Conspiracy Illusion	0.621	0.793	—	—
Tolerance for Ambiguity	0.611	0.345	0.824	—
Religious Orientation	0.601	0.104	0.573	0.781

4.1. Results

Figures 1 and 2 present the outcomes of hypothesis testing using the structural model. As shown in Figure 1, there was a statistically significant negative relationship between conspiracy illusion and auditor independence. The estimated path coefficient is −0.781, and the t-value is equal to 34.053. Being above the threshold, beyond 1.96 and establishing a lot of empirical support for the first research hypothesis

The result shown in Figure 2 indicates that the tolerance of ambiguity exerts a prominent moderating effect on the relationship between conspiracy illusion and auditor independence. A positive and significant path coefficient (0.501; t = 4.022) suggests that the relationship strengthens at lower values of ambiguity tolerance, and weakens for higher levels of this construct. In other words, among individuals who display greater tolerance for ambiguity, the adverse effect of conspiracy illusion on auditor independence is mitigated.

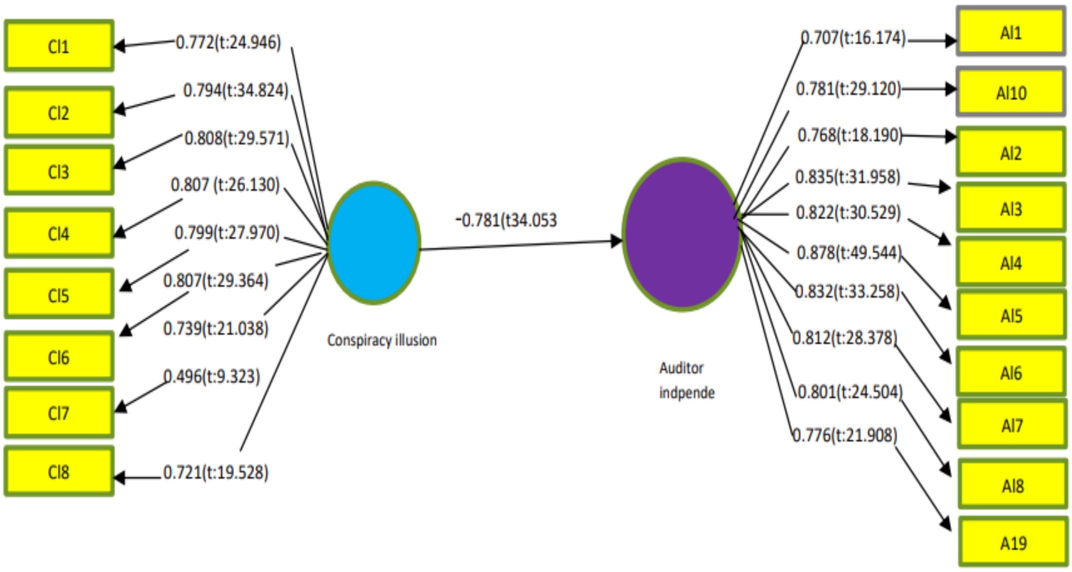

Figure 1. Results from the first hypothesis test

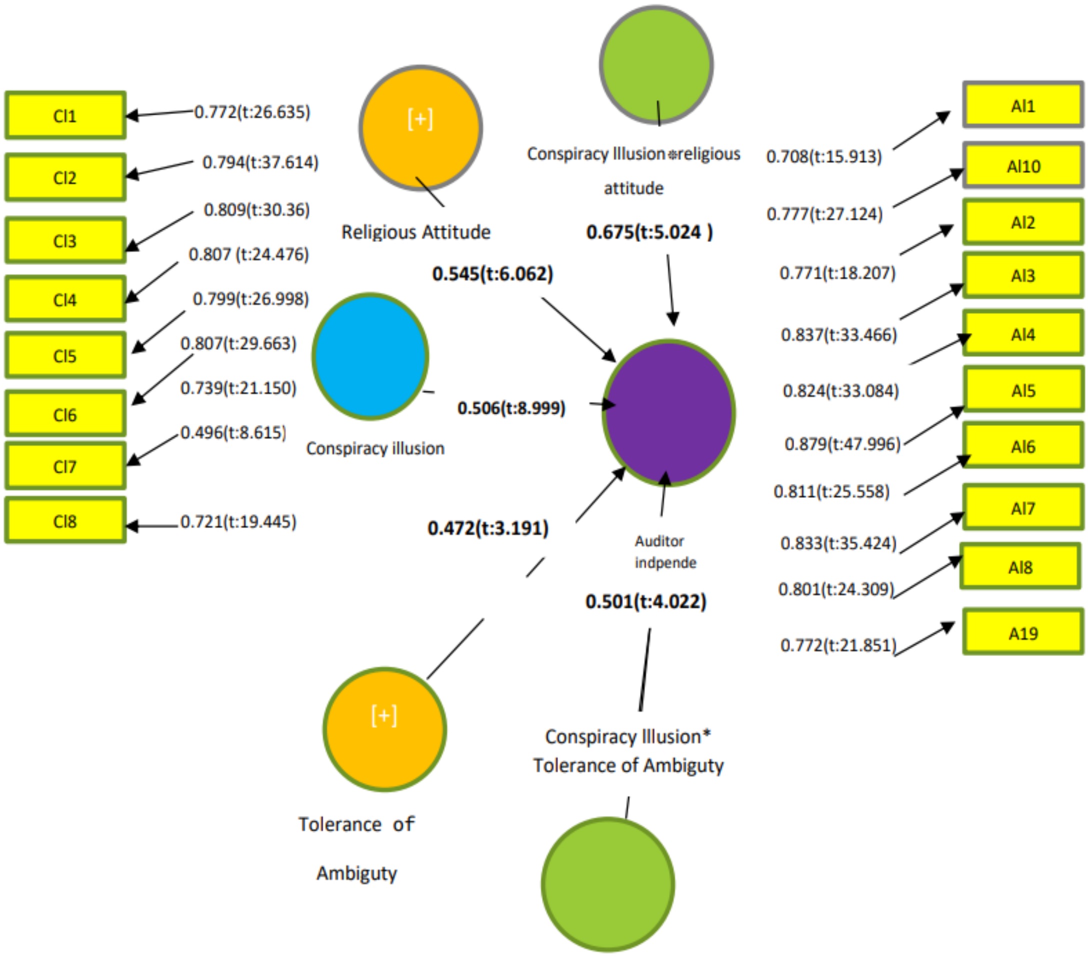

Figure 2. Results from the second hypothesis test

In addition, Figure 2 demonstrates that religious attitudes also play a significant moderating role in the relationship between conspiracy illusion and auditor independence. With a positive and statistically significant path coefficient (0.675; t = 5.024), the results suggest that stronger religious attitudes reduce the impact of conspiracy illusion on auditor independence. In other words, the interaction between the independent variable and the moderator exerts a diminished influence on auditor independence for individuals with higher levels of religiosity.

5. Conclusion

Auditor independence represents a cornerstone of audit quality, yet this study emphasized that independence extends beyond institutional or technical safeguards, which are deeply shaped by auditors' psychological characteristics. The findings showed that conspiracist ideation, a multidimensional cognitive bias, negatively impacted auditor independence. Auditors with higher levels of conspiracist ideation were more likely to engage in intuitive reasoning over analytical reasoning, especially when faced with ambiguous or stressful audit conditions. This cognitive style may have reinforced bias in evidence assessment, reduced appropriate levels of professional objectivity and fostered suspicion of organizational data and oversight mechanisms. These findings were in agreement with Elinda et al. (2019) that professional independence could be compromised because of psychological stress and pessimism and with Mardijuwono & Subianto (2018), who found excessive skepticism coupled with negative cognitive orientations to reduce judgment quality. In combination, these findings underscore that while professional skepticism is critical for sound auditing, the distorted version of such ideation as conspiracism can undermine, rather than support, independence. This relationship was moderated significantly by tolerance of ambiguity. Auditors who are more tolerant of ambiguity exhibited a greater psychological and cognitive resilience that enables them to cope with uncertain information better than their less tolerant counterparts, thus maintaining objectivity when faced with contradictory evidence. In contrast, individuals with low ambiguity tolerance were more prone to conspiratorial thought, which may have increased anxiety and eroded confidence in professional judgment. Likewise, religious orientation weakened the impact of a conspiracist ideation. Religious commitment was associated with less resistance to bias, stress or cognitive distortion among auditors when they hear these sounds. Religious orientation seemed to strengthen self-regulation, ethical reasoning, and value-based decision-making predictors that promote independence even when faced with pressure or uncertainty. In fact, this study demonstrated that auditor independence is dependent on statutory compliance and the technical capabilities of the audit partners and relies heavily upon their cognitive flexibility and ethical outlook. Conspiracist ideation endangered independence, whereas intolerance for ambiguity and religious orientation served to maintain psychological buffers. These results point out the need for both program and professional education in auditing to consider psychological development as a fundamental, integrating strategy alongside more traditional technical training. Efforts to increase auditors' ambiguity tolerance, ethical reasoning and self-awareness might also constitute effective means of bolstering independence and preserving public confidence in financial reporting.

Investigation to make sure there are training programs for auditors on how to identify and manage cognitive biases like conspiracist ideation, which identifies their effects on professional decision-making. Guided workshops with a focus on critical thinking, ambiguity resolution, and moral reasoning should systematically be included in professional development curricula for the strengthening of auditors' psychological resilience capacities and moral competence.

In addition, designing an organizational climate which is conducive to trust, as well as minimizing suspicion-based stories in audit teams, is useful in keeping conspiracist thinking at bay. Simultaneously, encouraging auditors to be more tolerant of ambiguity and hold a constructive religious orientation may help reduce the negative influence of conspiratorial cognition. Having auditors who display both social tolerance (for uncertainty) and ethical grounding may serve them in good stead to maintain professional independence, demonstrate balanced skepticism and make appropriate decisions under time pressure.

One of the main shortcomings of this study lies in the challenges that measuring such complex psychological constructs entails, such as measurement purposes (e.g., conspiratorial ideation, auditors' independence or ambiguity tolerance), ambivalence or religious orientation. Culture, context and respondents' subjective meanings may influence each of these variables, which can complicate the validity and comparability of results across different contexts (Dizon et al., 2020). Moreover, convenience samples of auditors may not reflect the entire auditing population and limit generalizability to broader audit contexts.

There is the possibility of social desirability bias in all self-reported data, where individuals answer based on what they believe others want to hear (e.g., how they ought to respond) rather than their lived experiences. Future studies should ideally address this limitation by including triangulation of data sources and methods (e.g., peer assessments, behavioral measures or archival audit performance records) to complement the presented subjective measures.

In practice, the results carry a few actionable implications. Audit regulators should seek to combine psychological competence assessments and ethics-oriented modules into professional certification and continuing education programs to ensure that auditors leave training equipped with the tools needed for recognizing cognitive biases, as well as for remediating. Large audit firms should create structured workshops, covering ambiguity management, critical thinking and ethical resilience, that enhance awareness of how psychological factors can compromise independence and judgment. Similarly, professional training institutions may integrate moral reasoning and emotional intelligence courses with bias awareness into auditor education curricula so that future professionals can maintain psychological stability and ethical integrity in tandem with technical proficiency.

Finally, qualitative approaches or experimental study designs would further help understand how these psychological and contextual variables affect auditor independence across cultures.

References

1. Abu Raiya, H., Pargament, K. I., Mahoney, A. and Stein, C. (2008). A psychological measure of Islamic religiousness: Development and evidence for reliability and validity. *The International Journal for the Psychology of Religion,* 18(4), pp. 291-315.
2. Azhdary, B., Nasl Mosavi, S. H. and Jafarilarijani, A. (2025). Investigating the effect of ethical principles, religiosity, auditor's personality and coronavirus on the quality of independent auditors' audit services. *Journal of Management Accounting and Auditing Knowledge,* 16(61), pp. 31-46 (In Persian). https://doi.org/10.22034/jmaak.2025.24065
3. Biddlestone, M., Green, R. and Douglas, K. M. (2020). Cultural orientation, power, belief in conspiracy theories, and intentions to reduce the spread of COVID-19. *British Journal of Social Psychology*, 59(3), pp. 663-673. https://doi.org/10.1111/bjso.12397
4. Citron, D. B. (2003). The UK's framework approach to auditor independence and the commercialization of the accounting profession. *Accounting, Auditing & Accountability Journal,* 16(2), pp. 244-274. https://doi.org/10.1108/09513570310472067
5. De Zavala, A. G., Cichocka, A., Eidelson, R. and Jayawickreme, N. (2009). Collective narcissism and its social consequences. *Journal of Personality and Social Psychology*, 97(6), A. 1074.
6. Douglas, K. M. (2021). Are conspiracy theories harmless?. *The Spanish Journal of Psychology*, 24, A. e13. https://doi.org/10.1017/SJP.2021.10

7. Douglas, K. M. and Sutton, R. M. (2018). Why conspiracy theories matter: A social psychological analysis. *European Review of Social Psychology,* 29(1), pp. 256-298. https://doi.org/10.1080/10463283.2018.1537428
8. Douglas, K. M., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S. and Deravi, F. (2019). Understanding conspiracy theories. *Political Psychology,* 40, pp. 3-35. https://doi.org/10.1111/pops.12568
9. Duh, R. R., Gul, F. A. and Hsu, A. W. H. (2022). Auditors' religious beliefs and audit quality: Some evidence from Taiwan. *The International Journal of Accounting,* 57(03), A. 2250013. https://doi.org/10.1142/S1094406022500135
10. Elinda, A. I., Iswati, S. and Setiawan, P. (2019). Analysis of the influence of role stress on reduced audit quality. *Jurnal Akuntansi,* 23(2), pp. 301-315. https://doi.org/10.24912/ja.v23i2.593
11. Ezzati Shalkoni, G. A., Jafari, S. M. and Darabi, R. (2024). The effect of the auditor's and employer's personality components on the development of religious attitude and the quality of audit. *Professional Auditing Research*, 4(14), pp. 56-87. https://doi.org/10.22034/jpar.2023.2004069.1173
12. Freeman, D., McManus, S., Brugha, T., Meltzer, H., Jenkins, R. and Bebbington, P. (2011). Concomitants of paranoia in the general population. *Psychological Medicine*, 41(5), pp. 923-936. https://doi.org/10.1017/S0033291710001546
13. Frenkel-Brunswik, E. (1949). Intolerance of ambiguity as an emotional and perceptual personality variable. *Journal of Personality,* 18(1), pp. 108–143.
14. Ghaznavi Doozandeh, J., Garkaz, M., Khozein, A. and Maetoofi, A. (2022). Identification of the factors affecting auditors' conflict of interests using fuzzy Delphi method. *Iranian Journal of Finance*, 6(1), pp. 142-168 (In Persian).
15. Golriz G. (1974). Study of the relationship between religious attitude and other feedback and job characterizations. B.Sc thesis in psychology, Faculty of Literature and Human Sciences, Tehran University, Tehran, Iran (In Persian).
16. Gunawan, D. and Lestari, M. A. (2025). Impact of auditor independence, professionalism, and skepticism on audit quality. *Advances in Accounting Innovation*, 1(2), pp. 178-187.
17. Jolley, D., Douglas, K. M., Skipper, Y., Thomas, E. and Cookson, D. (2021). Measuring adolescents' beliefs in conspiracy theories: Development and validation of the Adolescent Conspiracy Beliefs Questionnaire (ACBQ). *British Journal of Developmental Psychology*, 39(3), pp. 499-520. https://doi.org/10.1111/bjdp.12368
18. Karimi, Z., Yaghoob-Nejad, A., Samadi Largani, M. and Pourali Laklaye, M. R. (2021). Evaluating the impact of individual psychological biases and personality dimensions of auditors on audit quality. *Financial Accounting and Auditing Research*, 13(52), pp. 23–52 (In Persian).
19. Khaleq, S. (2023). The influence of cognitive bias on professional judgment in auditing: an empirical study from Egypt. *The Scientific Journal of Business and Environmental Studies*, 14(4), pp. 150-238.
20. Kurnia, R. (2021). The determinant factors affecting auditors' ability to detect fraud: Empirical study at public accounting firms in Jakarta and Tangerang. *Economics, Management and Sustainability*, 6(2), pp. 132-145.
21. Liu, S., Forgione, D., Xie, X. and Branson, L. (2014). Client importance and auditor independence. *Journal of International Business and Economics*, 14(4), pp. 95-110.
22. Mardijuwono, A. W. and Subianto, C. (2018). Independence, professionalism, professional skepticism: The relation toward the resulted audit quality. *Asian Journal of Accounting Research*, 3(1), pp. 61-71.

23. McLain, D. L. (2009). Evidence of the properties of an ambiguity tolerance measure: The multiple stimulus types ambiguity tolerance scale–II (MSTAT–II). *Psychological Reports*, 105(3), pp. 975-988. https://doi.org/10.2466/PR0.105.3.975-988
24. Sunstein, C. R. and Vermeule, A. (2009). Conspiracy theories: causes and cures. *Journal of Political Philosophy,* 17(2), pp. 202–227.
25. Swami, V., Chamorro-Premuzic, T. and Furnham, A. (2010). Unanswered questions: A preliminary investigation of personality and individual difference predictors of 9/11 conspiracist beliefs. *Applied cognitive psychology,* 24(6), pp. 749-761. https://doi.org/10.1002/acp.1583
26. Van Prooijen, J. W. and Jostmann, N. B. (2013). Belief in conspiracy theories: The influence of uncertainty and perceived morality. *European Journal of Social Psychology*, 43(1), pp. 109-115. https://doi.org/10.1002/ejsp.1922
27. Whitson, J. A. and Galinsky, A. D. (2008). Lacking control increases illusory pattern perception. *Science,* 322(5898), pp. 115-117.

IJAAF; Vol. 10, No. 2, Spring 2026, Serial 35

Ferdowsi University of Mashhad

Iranian Journal of
Accounting, Auditing & Finance

Quarterly

RESEARCH ARTICLE

A Comparative Study of XGBoost and Artificial Neural Networks for Earnings Management Prediction

Maedeh Moayeri*, Mohammad Arabmazar Yazdi, Vahid Menati

Department of Accounting, Faculty of Management and Accounting, Shahid Beheshti University, Tehran, Iran

How to cite this article:

Moayeri, M., Arabmar Yazdi, M. and Menati, V. (2026). A Comparative Study of XGBoost and Artificial Neural Networks for Earnings Management Prediction in Tehran Stock Exchange. Iranian Journal of Accounting, Auditing and Finance, 10(2), 69-87. doi: 10.22067/ijaaf.2026.47746.1579
https://ijaaf.um.ac.ir/article_47746.html

ARTICLE INFO

Article History
Received: 2025-08-09
Accepted: 2025-10-28
Published online: 2026-04-12

Keywords:
Artificial Neural Network (ANN), Earnings Management, Kasznik Model, Machine Learning, XGBoost

Abstract

The present study was conducted to compare the accuracy of ANN and XGBoost algorithms in predicting earnings management in listed companies of the Tehran Stock Exchange. Earnings management is one way for managers to mislead their stakeholders, which can result in financial losses; therefore, accurate detection methods are important for earnings management and beyond statistical models. The present study used 2016–2021 quarterly financial data from 103 publicly traded companies in basic metals, automotive, chemical, food and pharmaceutical producing industries (5076 year−firm). Discretionary accruals were used and calculated by the Kasznik model to capture earnings management and split them into three groups: Increasing Accruals (+1), Decreasing Accruals (-1), and Near-Zero Accruals (0). A Confusion Matrix was used to perform the model evaluation. The results showed that the XGBoost algorithm is significantly superior to the ANN with an overall accuracy of 98.4%. With very few errors, all earnings management categories XGBoost had near-perfect results. In contrast, ANN demonstrated significant weaknesses, leading to an overall accuracy of 63.1%. The study found that the model performance of XGBoost can further predict earnings management with more accuracy, thereby securing a process for financial institutions. Inspired by the relatively rare occurrence of applying the XGBoost model, used for trilateral classification of increasing, decreasing and near-zero earnings management in emerging markets, including the Tehran Stock Exchange. This research contributes significantly to the literature by demonstrating the superior predictive power of ensemble learning methods over traditional neural networks in detecting financial misconduct, which offers a robust, high-accuracy tool for regulators and investors to enhance market transparency and reduce financial risk.

https://doi.org/10.22067/ijaaf.2026.47746.1579

NUMBER OF REFERENCES
48

NUMBER OF FIGURES
4

NUMBER OF TABLES
5

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

*Corresponding Author: Maedeh Moayeri
Email: maedehmoayeri153@gmail.com
Tel: 09167381642
ORCID: 0009-0009-9308-9748

1. Introduction

Earnings management is essentially the manipulation of reported earnings accruals to achieve specific objectives (Saberi et al., 2023). In addition, earnings management occurs when managers use judgment in financial reporting and transaction structuring to alter financial reports, either to mislead some stakeholders about the company's underlying economic performance or to influence contractual outcomes that depend on reported accounting numbers (Healy and Wahlen, 1999). Understanding earnings management is possible in two ways: firstly, as opportunistic managerial behavior, it maximizes the apparent utility of management's performance for profit-based rewards; secondly, earnings management provides management with flexibility to use it for the benefit of related parties, in terms of anticipating unexpected events, for themselves and the company (Scott, 2015).

Accordingly, managers' motivations for earnings management can be divided into two distinct and opposing categories (Ahmadpoor and Montazeri, 2011):

1. **Efficient Earnings Management**: The objective of which is to enhance the quality of information preparation to help users better understand the company's profitability and financial position.
2. **Opportunistic Earnings Management**: Where managers undertake these actions solely to maximize their own interests, rather than those of investors.

If earnings management is efficient (good), earnings quality is expected to improve; conversely, if earnings management is opportunistic (bad), it may negatively impact earnings quality. Opportunistic behavior, by manipulating earnings, reduces the reliability of information and impairs earnings quality. The transmission of confidential information about future profitability opportunities to the market through intervention in the earnings measurement process affects earnings quality and consequently adds to the company's market value (Kordestani and Tatli, 2014).

In this research, opportunistic earnings management is of particular interest, as it can cause financial losses to financial statement users by creating misleading information. Therefore, detecting earnings management can prevent a significant portion of potential losses for users.

Accounting research over the years has proposed numerous models for identifying earnings management through financial statement items. These efforts began with seminal models such as the Healy Model (1985) and the DeAngelo Model (1996). The approach reached a turning point with the introduction of the Jones Model (1991) and subsequently the Modified Jones Model by Dechow et al. (1995), which attempted to control for non-discretionary factors. However, more sophisticated models, including the Kasznik Model (1999), were developed due to challenges arising from the insufficient control of firm operating performance. The Kasznik model enhanced the power to separate discretionary and non-discretionary components by adding the operating cash flow variable (Sun and Rath, 2010; Siekelova, 2021).

Accounting is the science of applying data, and in some cases, analyzing a large volume of data and making decisions based on it, which can be difficult and time-consuming. Given the increasing complexity of processes, the vast amount of information, and the intricate analysis required. Consequently, accounting researchers who need to automate these processes, like those in other fields, should explore the possibility of utilizing artificial intelligence tools. Machine learning is a subfield of artificial intelligence that focuses on developing algorithms capable of learning from data and making predictions or performing actions without explicit programming. The advantage of using these algorithms in accounting data analysis is saving time and analysis costs, and developing methods to increase the usefulness based on existing information and records. With the development of financial markets and the ever-increasing volume of information, participants in these markets are seeking simple and cost-effective tools that can enable them to make accurate and rapid predictions about the future market situation.

Although statistical models possess the ability to make acceptable predictions in analyzing financial and accounting issues, the limiting assumptions of some of these models (ambiguity in the relationships between input and output variables, collinearity error, lack of observations, etc.) reduce their effectiveness (Esfahanipour et al., 2016).

Moving beyond the established superiority of artificial intelligence methods over conventional statistical methods, we encounter a wide variety of machine learning algorithms. These algorithms themselves can exhibit varying levels of performance relative to each other, depending on their specifically defined objective. Therefore, researchers in this field are obligated to investigate two aspects: first, the feasibility of applying these algorithms to achieve the desired goal, and then to compare their accuracy and efficiency with each other to achieve the main objective of their research, which is to meet the needs of financial market participants. In this research, two powerful and novel machine learning algorithms were utilized (XGBoost and ANN) to predict earnings management in five industries: basic metals, automotive and parts manufacturing, chemical products, food products, and pharmaceutical materials and products, within the Tehran Stock Exchange market. We will then examine and compare their performance in terms of accuracy.

These two models, which are similar yet different in fashion in tackling complex financial data and the domestic literature, suffer from a critical research gap that makes the two of them the natural choices. The main advantage of ANNs is their inherent flexibility to model very non-linear relationships, which is a key requirement, considering that most financial data do not comply with simple linear structures (Artene and Domil, 2025). In contrast, XGBoost provides superior predictive accuracy, very fast run time, automatic handling of sparsity and robustness to different types of noise in the data compared with general linear models and many other conventional ensemble methods (Wiens et al., 2025). Importantly, most likely today, there is no evidence of any such literature that directly examines the comparative performance between XGBoost (as the leading gradient boosting algorithm) and Artificial Neural Network (ANN (a representative of non-linear/deep learning models) on predicting earnings management, in Iranian domestic literature.

Thus, the main purpose of this study is to set a new standard by directly comparing these two powerful algorithms under an adverse trilateral classification scenario (Increasing, Decreasing, and Near-Zero Accruals), with a precise Kasznik model. This comparison addresses a significant local research gap and provides a definitive assessment of the optimal predictive model for earnings management within the Tehran Stock Exchange.

The rest of this paper is organized as follows: We begin with Section 2, which establishes the theoretical foundations, reviews the relevant literature, and poses the research question. Next, Section 3 outlines the research methodology, including data collection and the implementation of machine learning algorithms. Subsequently, Section 4 reports the empirical findings and performance analysis, followed by the concluding remarks in Section 5.

2. Theoretical principles, literature review, and research question

2.1. Theoretical grounding of the prediction model structure

The structure of this prediction model is fundamentally based on the literature derived from Positive Accounting Theory (PAT) and Agency Theory. These frameworks provide the theoretical justification for selecting the financial features used as inputs (predictors) for the machine learning algorithms. The central theoretical premise is that managerial incentives (e.g., maximizing bonuses, avoiding debt covenant violations, mitigating political scrutiny) manifest as changes in financial performance indicators. The study uses the core components of the Kasznik Model, a theoretically justified accruals model, as the input features (X).

2.2. Earnings management

Like any other research field or domain, there is no consensus or agreement on a single, unified theory to explain and predict earnings management. The most debated theories of earnings management include: Prospect Theory, Catering Theory, Agency Theory, Signaling Theory, Big Bath Theory, Behavioral Theory, Game Theory, Shareholders Theory, and Entity Theory.

According to Saghafi and Pouryanasab (2010), all earnings management theories share three commonalities:

1. They all endeavor to explain the relationship between company managers and the group(s) of users or stakeholders concerning earnings management.
2. For this reason, and due to the use of two logical premises, "rational human" and "self-interest," and their product, "utility maximization," all these theories seek to explain the relationship between earnings management and value.
3. All these theories have been developed within the realm of positive accounting's dominance and therefore possess a positive and inductive nature, derived from observation and experience.

Ronen and Yaari (2002) accept "Firm Theory" as the dominant theory and place it above other theories. In this article, Firm Theory is also positioned as the dominant theory for three reasons: 1) It has a longer history than others, at least in other related research fields; 2) Other earnings management theories are, in some way, indebted to it, or in other words, this theory serves as a paradigm for the set of its rival theories; 3) And each of the other theories is, in some way, an approach derived from Firm Theory. As mentioned before, all earnings management theories strive to explain the relationship between insiders and outsiders regarding the phenomenon of earnings management. Therefore, a paradigmatic relationship exists between Firm Theory and its rivals, and thus, rival theories are each positioned within the framework of Firm Theory's approaches.

2.3. Models for measuring earnings management

Various models for detecting and measuring earnings management were proposed by different researchers, which are generally classified into two categories: discretionary accruals models, which are based on the Jones model, and discretionary revenue models, which have recently gained attention from financial researchers. Discretionary accruals, used as an indicator for measuring earnings management, are items over which management has control and can manipulate (accelerate or delay recognition, or omit them) (Salehi and Salehi, 2015).

Rahnamay Roodposhti et al. (2014) evaluated five accruals models and the Keeler discretionary revenue model for predicting earnings management in companies listed on the Tehran Stock Exchange. Their research findings indicated that earnings management models based on discretionary accruals have greater predictive power compared to the discretionary revenue model. Furthermore, the earnings management index derived from accrual-based models is a relevant and influential variable on the company's stock market value, in comparison to the revenue model.

To date, extensive research has been conducted in the field of comparing proposed models for detecting earnings management. Some of these studies in Iran have applied the proposed models in various industries and compared their results; for example, Montazeri and Khodamoradi (2017) investigated the detection power of earnings management models in companies listed on the Tehran Stock Exchange to identify the most suitable model for evaluating earnings management. In this research, the Jones, Modified Jones, and Kasznik models were compared. The findings indicated that among the three models, the Kasznik model yields the highest correlation. Also, according to the results of Rahmani and Bashirimanesh's research (2013), the adjusted Jones model was not significant in some industries and its adjusted coefficient of determination was also at a low level.

Kasznik model

The Kasznik model is as follows:

(Model 1)

$$\frac{TA_{it}}{A_{it-1}} = \alpha_0\left(\frac{1}{A_{it-1}}\right) + \beta_1\left(\frac{\Delta REV_{it} - \Delta REC_{it}}{A_{it-1}}\right) + \beta_2\left(\frac{PPE_{it}}{A_{it-1}}\right) + \beta_3\left(\frac{\Delta CFO_{it}}{A_{it-1}}\right) + \varepsilon_{it}$$

Where:

$TA_{i,t}$ = Accruals for company *i* in year *t*

$A_{i,t-1}$= Total book value of assets for company *i* at the beginning of the period

$\Delta REV_{i,t}$= Change in operating revenues for company *i* between year *t* and *t-1*

$\Delta REC_{i,t}$= Change in accounts receivable for company *i* between year *t* and *t-1*

$PPE_{i,t}$= Gross property, plant, and equipment for company *i* at the end of year *t*

ΔCFO_{it}= Change in operating cash flows between year t and t-1.

Kasznik's model adjusts the Jones model in three aspects. The first and most important adjustment is the inclusion of the change in cash from operations as the third independent variable (Salehi and Salehi, 2015). Recent empirical evidence and research show that accruals are negatively correlated with changes in cash flows, which is likely due to the nature of the accounting pattern (Dechow et al., 1995).

The second change from the Jones model is loosening the assumption of exogenous revenues. The timing of revenue recognition is a method often used by managers to manage reported earnings. If, in the model design, revenue is considered an exogenous factor beyond management's discretion, the presented model will not be able to detect earnings management through the timing of revenue recognition, and consequently, the model's power to detect earnings management will be weak (Salehi and Salehi, 2015). Following Dechow et al. (1995), this problem can be mitigated by adjusting the revenue variable by the change in accounts receivable.

The third adjustment involves the structure of the estimated portfolios. The accruals model uses a cross-sectional method instead of a time-series method to estimate the coefficients α_0, β_1, β_2, and β_3 (Salehi and Salehi, 2015). The advantage of the cross-sectional method is that it can control the effects of changes in economic conditions across a wide industry scope on total accruals and, due to the possibility of structural changes, allows for these coefficients to vary over the period (Kasznik, 1999).

2.4. Supervised machine learning

In machine learning methods, algorithms are developed that interpret the pattern of input data, extract the governing rules for that data set, and ultimately use these rules to predict other data.

Supervised learning is a type of machine learning algorithm where the model uses labeled training data and a set of training examples to infer a function (Sarker, 2020). Supervised learning occurs when we have identified and predetermined specific objectives for a particular set of data (Sarker et al., 2020). In essence, learning in this category of algorithms is used to identify patterns or behaviors in a "specified" training dataset (Boutaba et al., 2018).

2.5. Research background

First, relevant domestic research was reviewed to identify existing gaps in previous studies:

Faghani et al. (2016) aimed to predict earnings management using the Artificial Neural Network (ANN) and a hybrid Genetic Algorithm-ANN model. The findings showed that ANN has a high ability to predict earnings management compared to the linear Modified Jones model. However, this study's reliance on the Modified Jones model, which has been shown to have low detection power in certain industries, limits its generalizability (Montazeri and Khodamoradi, 2017).

Salehi and Farokhi Pile Rood (2018) compared the predictive accuracy of earnings management using Neural Networks and Decision Trees against linear models. The results indicated that neural networks and decision trees offer greater accuracy and lower error rates in predicting earnings management compared to linear methods. Given the established superiority of AI tools over traditional linear models, the current research moves beyond this conclusion to find the single best-performing algorithm.

Fereydouni et al. (2020) predicted earnings smoothing using the Relevance Vector Machine (RVM) algorithm. The results showed that the RVM algorithm, in both its linear and non-linear forms, has a high capability in predicting the extent of earnings smoothing, with the non-linear approach performing better. The current research contributes by comparing and examining newer and more powerful machine learning algorithms to achieve a more comprehensive conclusion regarding the optimal model for earnings management prediction.

Maleki Kakler et al. (2021) applied machine learning models (Decision Tree, Support Vector Machine, and Bayesian Network) to develop an effective pattern for predicting fraudulent financial reporting. The results showed that the Decision Tree algorithm was successful, achieving an accuracy of over 92.61%. Similar to others, this study focused on demonstrating the superiority of machine learning over statistical models, whereas our research centers on a head-to-head comparison of two distinct and powerful machine learning algorithms.

Hassani et al. (2024) conducted a study to distinguish between accrual earnings management and real earnings management using machine learning methods (decision trees, SVM, KNN, deep learning) and combining them with relief feature selection and principal component analysis (PCA) methods. The findings showed that in predicting accrual earnings management, the relief-based feature selection model has a better ability than the PCA model, but this superiority was not observed in predicting real earnings management. Also, the results indicated that accrual earnings management can be predicted with higher accuracy than real earnings management. In this study, a binary classification method was used, while the present study uses a triple classification method.

Hamidian and Esmaeili (2024) have a study focused on predicting earnings management using support vector machine and decision tree methods in 170 companies. The results showed that support vector machine and decision tree methods have the ability to predict earnings management of companies and the support vector machine method was introduced as the superior method in predicting earnings management. In this study, the focus is more on the independent variables (attributes) than on the specific model of measurement of the dependent variable (discretionary accruals).

After mentioning the gaps in domestic research in the field of applying machine learning in earnings management prediction, studies that are more closely related to the research objective and methodology can be reviewed, focusing on finding new research opportunities:

Rahman et al. (2021) compared two types of machine learning algorithms, linear classification and tree-based classification (Decision Tree and Random Forest), for predicting earnings manipulation. The results indicated that the performance of the two algorithm types was not significantly different, but tree-based classifiers achieved the best accuracy in the shortest time.

Hammami and Hendijani Zadeh (2022) employed six new data analysis methods in classification to predict earnings management (both accrual and real). The results showed that the ABC-SVM model (Ant Colony Optimization and Support Vector Machine) performed better than other models in predicting earnings management.

Chen et al. (2022) combined Random Forest with Gradient Boosting in a hybrid machine learning model to predict one-year-ahead earnings changes in direction. The accuracy of the proposed model was found to be between 67.52% and 68.66%, which outperforms the linear models.

Adomat et al. (2023) investigated a new cloud production environment-based earnings management technique using machine-learning algorithms. They observed that their combined Neural Network and Linear Regression model achieved the lowest error rates compared to standard and automated linear regression algorithms.

Ali et al. (2023) produced a state-of-the-art FSF detection model dedicated to the MENA region with publicly available financial data. Based on the empirical results, XGBoost performed substantially higher than all other machine learning models - Logistic Regression, Decision Tree, Support Vector Machine, AdaBoost and Random Forest. Optimization resulted in a final accuracy of 96.05% for FSF detection by the XGBoost algorithm.

Chakri et al. (2023) used four different supervised machine learning models (Linear Regression, K-Nearest Neighbors, Support Vector Regression, and Decision Tree) for earnings prediction. Decision Tree proved to be the best model for performance analysis according to the findings.

In Chen et al. (2025), the multi-class Financial Distress Prediction (FDP) model was developed, which offered a deeper insight into corporate financial distress classification. The study introduced a new definition of the financial status in three classes and applied the Deep Forest algorithm for classification. This is the first multi-class forecast in the financial domain, and thus is a remarkable research.

Huy et al. (2025) addressed the limitations of traditional discretionary accrual models in detecting the nonlinear patterns of earnings management. They evaluated tree-based machine learning models (Decision Trees, Random Forests, and Gradient Boosting Machines). Their findings established that the Gradient Boosting Machine (GBM) significantly outperformed the other models across all key metrics (accuracy, precision, recall, and F1 score), proving to be the most effective tool.

The aim of Veganzones and Severin (2025) was to identify the financial profiles of companies using machine learning and to provide a visual representation of the financial profiles associated with earnings management strategies (upward and downward) and tools (accruals and real activities). The proposed machine learning method performed better than traditional forecasting methods.

Domestic literature on predicting earnings management using machine learning algorithms in the Tehran Stock Exchange is constrained by a reliance on older models (proving only the general superiority of Machine Learning over linear methods), the use of less accurate earnings management models like the Modified Jones model (Montazeri and Khodamoradi, 2017; Rahmani and Bashirimanesh, 2013) and a simplistic binary classification.

This study addresses these limitations by establishing a new benchmark for earnings management prediction on the Tehran Stock Exchange. Our innovation lies in three key areas: 1) Employing the Kasznik model, which is locally validated as more robust; 2) Adopting a challenging trilateral classification (increasing, decreasing, and near-zero earnings management); and 3) Conducting a systematic, head-to-head comparison between the XGBoost and Artificial Neural Network (ANN) algorithms to identify the optimal model in this field.

2.6. Research question

Which of the Artificial Neural Network (ANN) algorithms and XGBoost has higher accuracy in predicting earnings management?

3. Research methodology

The present research, in terms of its objective, is applied; and in terms of its nature, given that it examines and compares the accuracy of algorithms and uses real data for testing, it falls under descriptive research. Regarding the research method, an ex-post facto approach was used. In ex-post facto research, the investigation begins after an event has occurred. Here, the researcher, essentially not present, does not manipulate the variables and does not know them, but rather conducts causal research to identify these variables and factors that caused the event (Sarmad et al., 2004). In this research, by implementing machine learning algorithms, we aim to improve earnings management prediction and compare the results of the algorithms with each other to ultimately select the best one (in terms of accuracy) for better prediction.

3.1. Data collection method

Many accounting and financial studies on the Tehran Stock Exchange are conducted using the Rahavard Novin database. Rahavard Novin is the most comprehensive software for the Iranian capital market. Since its initial release in 1999, due to its extensive database and advanced tools, it has always been the primary choice for both individual and institutional participants in the capital market (Mosallanejad and Nazemi, 2019). Rahavard Novin software provides advanced tools that allow access to extensive information on companies listed on the Tehran Stock Exchange in an optimal time. In this research, a complete set of required information was extracted from this software, and when necessary, information was verified by direct reference to the Codal website.

3.2. Statistical population and sample

The statistical population of this research consists of classified and audited financial data of active companies listed on the Tehran Stock Exchange during the period from 2016 to 2021.

The statistical sample (present in Table 1) is determined by the systematic exclusion method and includes companies that meet the following conditions:

1. Must be a listed company on the stock exchange.
2. The company's fiscal year must end on March 20.
3. The company's fiscal year must not have changed within the research period (2016 to 2021).
4. Must have been continuously active on the stock exchange from 2016 to 2021.
5. Required financial information, including financial statements and accompanying notes, for calculating research variables must be available for the relevant companies within the 2016-2021 timeframe.
6. Must belong to one of the five industries: Basic Metals, Automobile and Parts Manufacturing, Chemical Products, Food Products, and Pharmaceutical Materials and Products.

The table related to the application of each of the above-mentioned restrictions is as follows:

Table 1. Sample selection process by applying population constraints and conditions

Constraint	Percentage of Deletion	Remained Sample Size
Initial population		**5076 year-firm**

1. Companies must be listed on the main stock exchange (Tehran Stock Exchange) and not on the over-the-counter market.	≅ 62%	1892 year-firm
2. The companies' fiscal year must end on March 20.	≅ 45%	1032 year-firm
3. The companies' fiscal year must not have changed within the research period.	≅ 14%	880 year-firm
4. Must have been continuously active on the stock exchange from 2016 to 2021.	≅ 5%	836 year-firm
5. Required financial information, including financial statements and accompanying notes, for calculating research variables must be available for the relevant companies within the 2016-2021 timeframe.	≅ 2%	816 year-firm
6. Must belong to one of the five industries: Basic Metals, Automobile and Parts Manufacturing, Chemical Products, Food Products, and Pharmaceutical Materials and Products.	≅ 49%	412 year-firm (Final Sample)

Note. The total number of companies at the beginning of the selection process was 5076 year-firm. The percentages indicate the estimated proportion of deletion at each step relative to the previous sample size.

3.3. Obtaining discretionary accruals

In this stage of the research, the variables of interest (according to the Kasznik model), including the book value of company assets, operating income, net accounts receivable, and operating cash flow, were extracted from the available data in the sample's financial statements and saved in an Excel environment. Gross property, plant, and equipment (PPE) is among the information found in the notes to the financial statements, which was collected by direct reference to the financial statements via the Codal website.

In the Kasznik model, β and α are company-specific parameters estimated using OLS regression.

Given the total accruals (TA), which comprise normal accruals (NA) and abnormal accruals (AA), the sum of the four terms on the right side of the Kasznik model (Model 1) represents normal accruals, and $\varepsilon_{i,t}$ (the residual) indicates the abnormal (discretionary) accruals for company *i* in fiscal year *t*. Discretionary accruals are recognized as a measure of earnings management, and their level indicates the occurrence or non-occurrence of earnings management (Dechow et al., 1995).

According to Guara et al., total accruals (TA) are calculated using the following formula:

(Formula 1)

Total Accruals = (Net Income - Operating Cash Flow) × (-1)

Accordingly, the level of discretionary (abnormal) accruals is measured after estimating the regression coefficients (β and α) for each company separately using OLS regression and measuring ε_{it}.

3.4. The role of the Kasznik model in the machine learning framework & data labeling

The Kasznik Model (1999) plays a foundational and dual role in establishing the Supervised Machine Learning framework of this study, which serves as both the measurement tool for the target variable (Y) and the basis for defining the input features (X).

3.4.1. Defining the dependent variable (Y) and labeling

Measurement Tool (Target Definition): The discretionary accruals (ϵi,t) obtained in Section 3-3, which serve as the proxy for earnings management, are the core input for data labeling. Data labeling is performed based on the value obtained from the level of discretionary accruals. Based on predefined thresholds, this value is converted into the discrete categorical variable Y (the dependent variable for prediction), classifying the outcome into Increasing Accruals (+1), Decreasing Accruals (-1), or Near-Zero Accruals (0).

3.4.2. Defining the input features (X) and timeliness control

Feature Definition: The model also fundamentally dictates the selection of our input features. The input variables, directly extracted from the sample companies' financial statements (Balance Sheet, Income Statement, and Cash Flow Statement) at the end of year t−1, correspond to the components of the Kasznik Model (1999), which are the standard determinants used to estimate Non-Discretionary Accruals.

The machine learning models (XGBoost and ANN) were trained to predict earnings management in year t. Accordingly, all input variables (Features) were exclusively extracted from financial information available at the end of the fiscal year t−1 (lagged variables). This rigorous time control ensures that the models use no information that would not have been published and available at the time of the actual prediction (i.e., at the beginning of year t), thereby eliminating the risk of Look-ahead Bias and Data Leakage.

3.5. Examining the general distribution of discretionary accruals

Figure 1 is a histogram illustrating the distribution of data related to discretionary accruals (earnings management) within the examined dataset. A histogram is a type of bar chart for visualizing continuous data, which shows the distribution of values within a dataset (Archambault et al., 2015). In this graph, a probability distribution curve has also been used to more clearly illustrate the data distribution. This graph indicates that the highest data density (1.42) is at a point close to zero (-0.04). According to this graph, the data distribution is asymmetric, non-normal, and has a long tail to the right, which suggests that high values of earnings management are more frequent than low values.

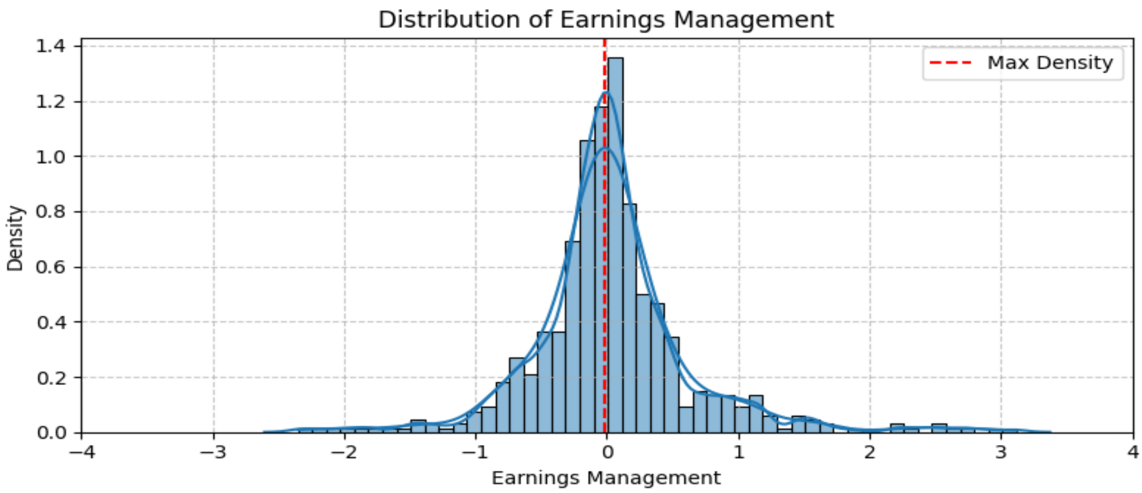

Figure 1. Distribution of data related to discretionary accruals (earnings management)

3.5.1. Determining the first and third quartiles and the median of the data

In this distribution, the median, first quartile, and third quartile values are as follows (Figure 2):

- *Median:* 0.0018487886711916826
- *First Quartile (Q1):* -0.22011942451641645
- *Third Quartile (Q3):* 0.23925656442091495

Based on the obtained values:

- Discretionary accruals whose value is greater than or equal to the third quartile are classified as items with increasing earnings management and labeled as +1.

- Discretionary accruals whose value is less than or equal to the first quartile are classified as items with decreasing earnings management and labeled as -1.
- Discretionary accruals whose value is between the first and third quartiles are classified as items with near-zero earnings management and labeled as 0.

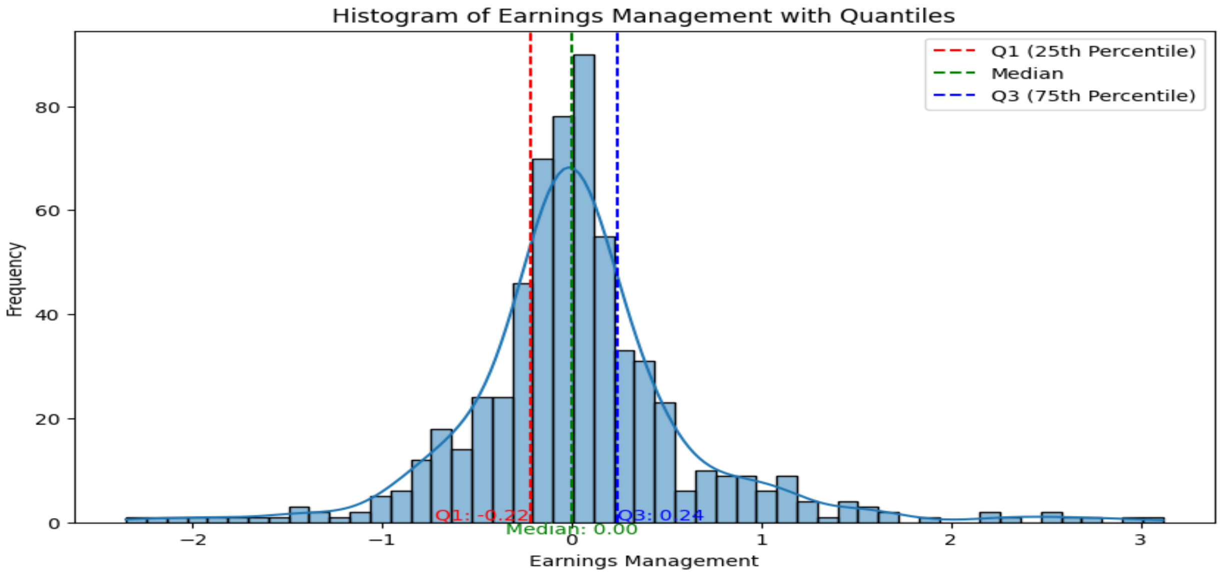

Figure 2. First quartile, third quartile, and median of the discretionary accruals data distribution

3.5.2. Splitting training data into training and test sets

An algorithmic model is built based on training data, and then the model's performance is evaluated using test data (Jain et al., 2022). In fact, at this stage of the research, we focus on developing algorithmic models and adapting them to the data structure, which is considered the main task in this research.

Overfitting is a fundamental problem in supervised machine learning that prevents fully generalizing the model from observed training data to unobserved test data and can arise due to noise, the small size of the training set, and the complexity of the classifier. Despite overfitting (the model works perfectly on the training set and suboptimally on test data), it has some issues dealing with information that is more present in the test set than what was seen during training (Ying, 2019). In order to minimize the probability of overfitting, we take some part of the training dataset as a validation set.

This process is called the hold-out method. We take a label dataset (training) and split it into a training and test set. Next, we train a model on the training set and make predictions for the test set labels. We compute the fraction of correct predictions (via comparison to the pre-predicted labels), which is our estimate of predictive accuracy for the model (Burzykowski et al., 2023).

Splitting the training dataset into training and test sets is considered essential for eliminating or reducing bias in training data within machine learning models. This process is always performed by data scientists and data analysts to prevent algorithm overfitting (Muraina, 2022).

The hold-out method is one of the most common cross-validation methods. The set of cross-validation techniques is a method used to evaluate the predictive performance of a model (Yadav and Shukla, 2016).

This operation is performed on the training set with the following code example:

```
# Split the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split (X, y, test_size=0.2, random_state=42)
```

The percentage for splitting training and test data is not considered significant in any source and can be chosen by the researcher based on the total volume of training data. In this research, 80% of the training dataset is used for model fitting, and 20% of this set is considered for evaluating the model's accuracy (test data). This splitting process was performed using the stratified hold-out method to ensure that the proportional representation of the three earnings management classes is maintained across both the training and test sets, thereby mitigating bias arising from class imbalance.

3.6. Implementation of machine learning algorithms

In this section, the machine learning models of interest for this research, including K-Nearest Neighbors, Support Vector Machine, Decision Tree, and Random Forest, are each fitted separately and their performance evaluated.

Model fitting is performed by estimating the most suitable hyperparameters. Hyperparameter estimation is carried out using GridSearch. This method simply performs an exhaustive search over a specified subset of the training algorithm's hyperparameter space.

Since the hyperparameter space of a machine learning algorithm may include a space with infinite values for some parameters, we must specify a boundary for each hyperparameter to apply Grid Search (Liashchynskyi and Liashchynskyi, 2019). To mitigate the instability of a single Hold-out split and ensure robustness and generalization capability, GridSearch inherently employs k-fold cross-validation (CV) (specifically, we utilized 5-fold cross-validation here) on the training set to select the optimal parameters

XGBOOST Algorithm XGBoost is a machine learning algorithm that utilizes ensemble tree models. Inspired by tree boosting algorithms (composed of decision trees), it uses a combination of multiple Regression Trees as its core model. In this method, each leaf of the tree contains a continuous score. Decision rules within the trees are used to classify these leaves. The final class of each leaf is predicted by aggregating the scores of the corresponding leaves, striving to provide the best prediction for machine learning problems (Rashidpour and Zare Bidaki, 2024).

The XGBoost algorithm has numerous hyperparameters. To achieve the best predictive performance and ensure robustness, the optimal values for the most critical hyperparameters were determined using Grid Search combined with 5-fold Cross-Validation. The three key hyperparameters selected for tuning were the n_estimators (number of decision trees), max_depth (maximum depth of decision trees), and learning_rate (shrinkage rate). Their optimal values are summarized in Table 2.

Table 2. XGBoost best Hyperparameters

Hyperparameter	Value
$N_{estimator}$	50.000
Learning rate	0.010
Max_{depth}	3.000

ANN Algorithm, an Artificial Neural Network (ANN), is a type of mathematical model that mimics the functioning of the human brain. Their ability lies in extracting patterns from observed data without requiring assumptions about the relationships between variables. They are comprehensive, flexible functions and powerful tools for data analysis and modeling nonlinear relationships with a high degree of accuracy (Saberi et al., 2023).

ANN encompass various types of models, and in this research, we utilize the Deep Neural Network model. A Deep Neural Network is a multi-layered neural network that possesses several hidden layers. This algorithm also includes hyperparameters, whose optimal values were obtained using GridSearch and are presented in Table 3.

Table 3. ANN best Hyperparameters

Hyperparameter	**Value**
$Batch_{size}$	10
epochs	50
$Model_{neurons}$	32
$Model_{optimizer}$	adam

3.7. Programming environment and software implementation

The implementation, training, and testing of the machine learning algorithms (XGBoost and ANN) were performed using the Python 3.9 programming language within the Visual Studio Code (VS Code) environment. The following specialized libraries were utilized for data processing and model execution:

- *Core Libraries:* Scikit-learn (for model implementation, cross-validation, and data splitting), XGBoost (for the Extreme Gradient Boosting model), TensorFlow/Keras (for building and training the Artificial Neural Network), Pandas and NumPy (for data manipulation and algebraic operations).
- *Statistical Libraries:* Statsmodels (utilized for the OLS regression necessary to estimate the Non-Discretionary Accruals in the Kasznik model).

This detailed documentation ensures the reproducibility of our research findings.

4. Research findings

To evaluate the performance of the models, we utilize the Confusion Matrix. A Confusion Matrix is a table that visualizes the performance of a supervised learning algorithm. Each column of the matrix represents the predicted classes, while each row represents the actual classes. All correct predictions are located along the diagonal of the matrix, allowing errors to be easily observed by non-zero values outside the diagonal (Patro and Patra, 2014).

Precision, Recall, F1-score, and Support are metrics extracted from the Confusion Matrix to assess algorithm performance:

- *Precision:* Among all instances that the model predicted to belong to a specific class, what percentage truly belonged to that class? High precision means a low False Positive rate.
- *Recall:* Among all instances that actually belonged to a specific class, what percentage did the model correctly identify? High recall means a low false-negative rate.
- *F1-score:* Represents the harmonic mean of precision and recall. A high F1-score indicates good performance in both precision and recall.
- *Support:* Denotes the actual number of occurrences for each class in the specified dataset.
- *Weighted Average:* reports the model's performance across all classes (Categories) simultaneously, by taking into account the size (Support) of each individual class.

For the XGBoost algorithm, the results, according to the Confusion Matrix, are as follows (Figure 3 and Table 4):

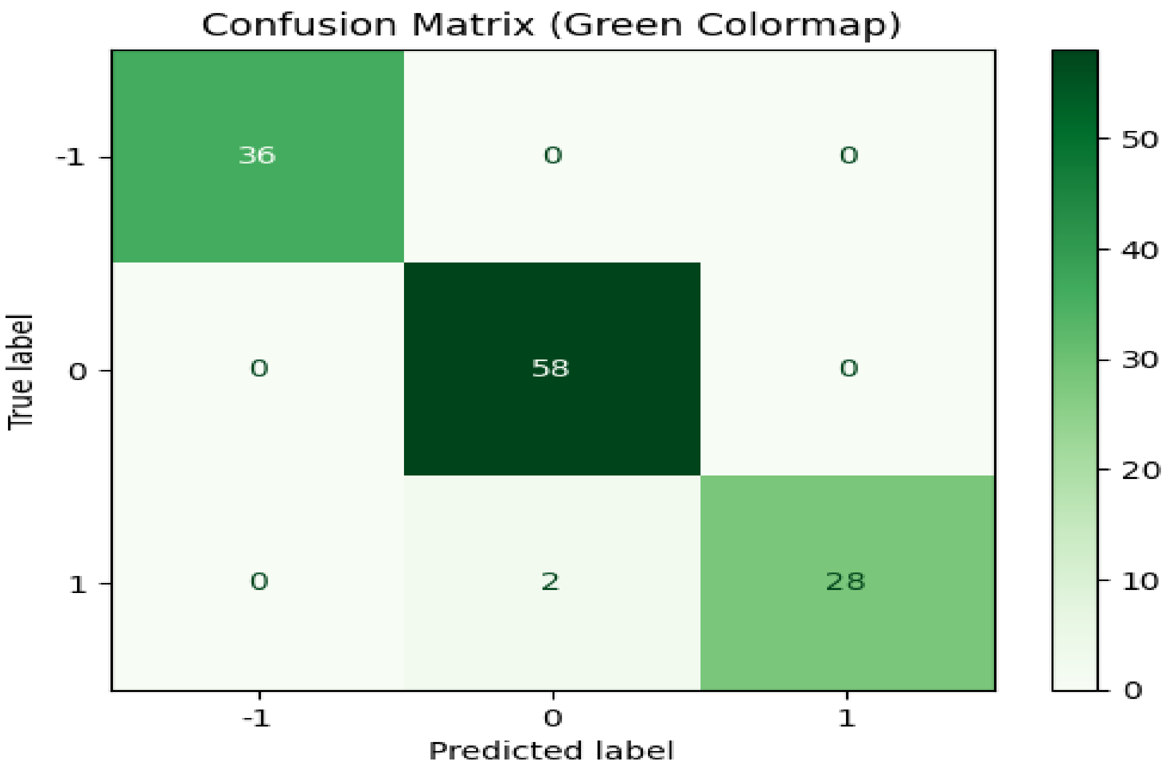

Figure 3. Confusion matrix of the XGBoost algorithm

Table 4. Performance metrics for the XGBoost algorithm (Classification Report)

Class	Precision	Recall	F1-Score	support
Decreasing Accruals (-1)	1.000	1.000	1.000	36
Near-Zero Accruals (0)	0.970	1.000	0.980	58
Increasing Accruals (+1)	1.000	0.930	0.970	30
Weighted Average	0.980	0.980	0.980	124

Note. The reported metrics are based on the testing set (N=124). The overall model accuracy was 0.98.

4.1. Interpretation of XGBoost algorithm results

This model demonstrated perfect performance for Class -1, which correctly identified all 36 instances of Class -1; every time it predicted an instance as Class -1, it was correct.

Furthermore, in identifying Class 0, it correctly found all 58 instances (recall: 1.00). When it predicted an instance as Class 0, it was correct 97% of the time, meaning there were very few false positives for this class (precision: 0.97).

Finally, every time it predicted Class 1, it was correct (precision: 1.00). However, its Recall is slightly lower at 93%, meaning it did not identify 7% of the actual Class 1 instances (i.e., it had false negatives for Class 1).

Similarly, for the ANN algorithm, the results according to the Confusion Matrix are as follows (Figure 4 and Table 5):

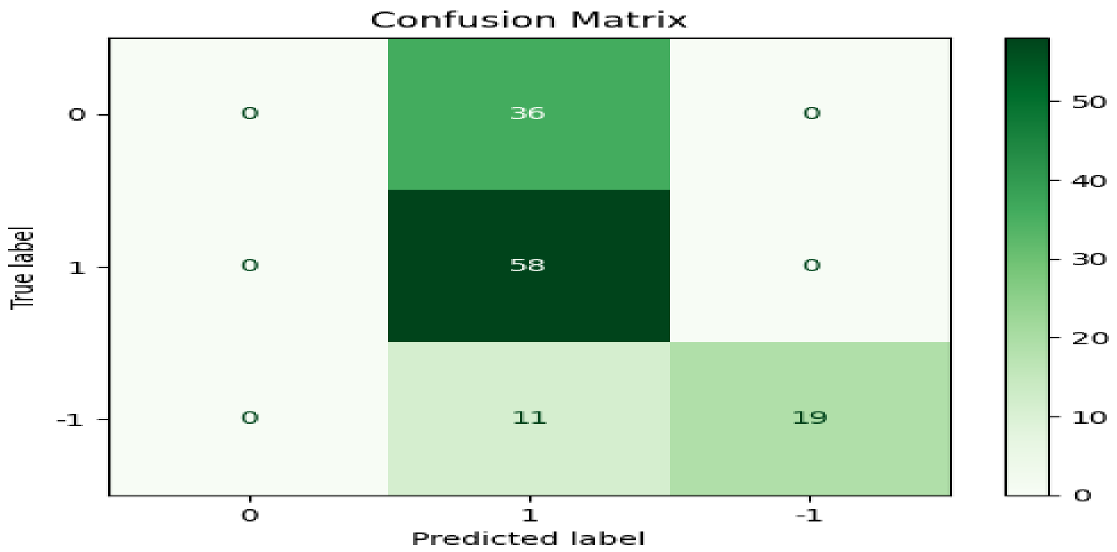

Figure 4. Confusion matrix of the ANN Algorithm

Table 5. Performance metrics for the ANN algorithm (Classification Report)

Class	Precision	Recall	F-Score	support
Decreasing Accruals (-1)	0.000	0.010	0.020	36
Near-Zero Accruals (0)	0.560	1.000	0.720	58
Increasing Accruals (+1)	1.000	0.670	0.800	30
Weighted Average	0.630	0.630	0.630	124

Note. The reported metrics are based on the testing set (N=124). The overall model accuracy was 0.63.

4.2. Interpretation of ANN algorithm results

Recall (Sensitivity): All actual Class 0 instances were correctly classified. Precision: Only 55% of the instances predicted as Class 0 by the model were actually Class 0.

All predicted Class 1 instances were actually Class 1. This means that the model has a very high precision in predicting for this class (Precision: 1.00). The recall for Class 1 was at only 63%. That is to say, the model classified 37% of true Class 1 cases as another class (Recall: 0.63).

For Class -1, since there were no instances that belonged to it (highlighted in red, middle column — by default), the Precision and Recall are 0.

5. Conclusion

The primary purpose of this research was to compare the performance of the Artificial Neural Network (ANN) and XGBoost algorithms in predicting earnings management among the companies of five specified industries) listed on the Tehran Stock Exchange. The novelty of this study lies in its use of a trilateral classification of earnings management (Increasing Accruals: +1, Decreasing Accruals: -1, and Near-Zero Accruals: 0) and the application of the Kasznik Model for predicting.

5.1. Methodology and key findings

This study utilized financial data from 103 companies (5076 year-firm) across five industries (basic metals, automotive, chemical, food, and pharmaceutical) over the period 2016 to 2021. The research adopted an applied, descriptive, ex-post facto approach. Discretionary accruals, derived from the Kasznik model, served as the proxy for earnings management, and the data were strictly labeled into the three classes based on quartiles. To gain a valid prediction of current-year (t) earnings management and to thwart implicit data leakage, a crucial methodological control was instituted by limiting all input variables to information available at the end of the prior fiscal year (t-1). A confusion matrix was used to evaluate the performance of the models.

The findings above clearly demonstrated that the XGBoost algorithm outperforms:

- XGBoost Algorithm: 98.4% overall accuracy demonstrated perfect or near-perfect performance across all three earnings management categories, successfully differentiating between them with very few errors.
- ANN Algorithm: Showed considerable weaknesses, resulting in a significantly lower overall accuracy of 63.1%. The ANN model was completely unsuccessful in identifying the Near-Zero Accruals (Class 0), demonstrating a failure to generalize across all classes.

The study concludes that XGBoost is a superior and more reliable model for predicting earnings management in this context.

5.2. Comparison with past research

This research contributes significantly to the literature by moving beyond the established superiority of machine learning methods over conventional statistical models. Previous studies on the Tehran Stock Exchange have often been limited by using older earnings management models like the Modified Jones model (which has low detection power in some industries) and simpler binary classification. Leveraging the advanced Kasznik model and a complex 3-way classification, this study lays the groundwork for future research and demonstrates that set-based methods, such as XGBoost, surpass traditional neural networks (ANN) in addressing sophisticated financial misbehavior.

5.3. Practical implications and future research suggestions

Practically, the high-accuracy XGBoost model offers a robust tool for regulators and investors to enhance market transparency and reduce financial risk by identifying companies engaging in opportunistic earnings management.

For future research, it is suggested:

1. To extend the application of the XGBoost model to predict other forms of financial misconduct, such as fraudulent financial reporting, and compare its efficacy with other advanced ensemble or deep learning algorithms.
2. To replicate this trilateral classification methodology in other industries to test the generalizability of XGBoost's superior performance.

5.4. Research limitations

The conclusions of this study are limited to the scope of its design; Specifically:

- The findings are restricted to the five specified industries and companies listed on the Tehran Stock Exchange.
- The superior effectiveness of the XGBoost model is only established against the Artificial Neural Network within the context of trilateral classification.
- It is important to note that this study's conclusions focus solely on predictive superiority and do not extend to causal analysis of earnings management.

References

1. Adomat, V., Ehrhardt, J., Kober, C., Ahanpanjeh, M. and Wulfsberg, J. P. (2023). A machine learning approach for revenue management in cloud manufacturing. *Procedia CIRP,* 118, pp. 342-347. https://doi.org/10.1016/j.procir.2023.06.059
2. Ahmadpoor, A. and Montazeri, H. (2011). Type of earnings management and the impact of company size, ownership structure and corporate governance on earnings management. *Journal of Accounting Advances*, 3(2), pp. 1-35. https://doi.org/10.22099/jaa.2012
3. Ali, A. A., Khedr, A. M., El-Bannany, M. and Kanakkayil, S. (2023). A powerful predicting model for financial statement fraud based on optimized XGBoost ensemble learning technique. *Applied Sciences*, 13(4), A. 2272. https://doi.org/10.3390/app13042272
4. Artene, A. E. and Domil, A. E. (2025). Neural networks in accounting: bridging financial forecasting and decision support systems. *Electronics,* 14(5), p. 993. https://doi.org/10.3390/electronics14050993
5. Boutaba, R., Salahuddin, M. A., Limam, N., Ayoubi, S., Shahriar, N., Estrada-Solano, F. and Caicedo, O. M. (2018). A comprehensive survey on machine learning for networking: evolution, applications and research opportunities. *Journal of Internet Services and Applications*, 9(1), pp. 1-99. https://doi.org/10.1186/s13174-018-0087-2
6. Burzykowski, T., Geubbelmans, M., Rousseau, A. J. and Valkenborg, D. (2023). Validation of machine learning algorithms. *American Journal of Orthodontics and Dentofacial Orthopedics*, 164(2), pp. 295-297. https://doi.org/10.1016/j.ajodo.2023.05.007
7. Chakri, P., Pratap, S. and Gouda, S. K. (2023). An exploratory data analysis approach for analyzing financial accounting data using machine learning. *Decision Analytics Journal*, 7(1), A. 100212. https://doi.org/10.1016/j.dajour.2023.100212
8. Chen, X., Cho, Y. H., Dou, Y. and Lev, B. (2022). Predicting future earnings changes using machine learning and detailed financial data. *Journal of Accounting Research,* 60(2), pp. 467-515. https://doi.org/10.1111/1475-679X.12429
9. Chen, X., Mao, Z. and Wu, C. (2025). Multi-class financial distress prediction based on feature selection and deep forest algorithm. *Computational Economics*, 66(4), pp. 2715-2754. https://doi.org/10.1007/s10614-024-10761-8
10. DeAngelo, H., DeAngelo, L. and Skinner, D. J. (1996). Reversal of fortune dividend signaling and the disappearance of sustained earnings growth. *Journal of Financial Economics,* 40(3), 341-371. https://doi.org/10.1016/0304-405X(95)00850-E
11. Dechow, P. M., Sloan, R. G. and Sweeney, A. P. (1995). Detecting earnings management. *Accounting Review*, 70(2), pp. 193-225.
12. Esfahanipour, A., Goodarzi, M. and Jahanbin, R. (2016). Analysis and forecasting of IPO underpricing. *Neural Computing and Applications,* 27(3), pp. 651-658. https://doi.org/10.1007/s00521-015-1884-1
13. Faghani Makrani, K., Salehnejad, S. H. and Amin, V. (2016). Predicting earnings management based on modified Jones model using artificial neural network model and genetic algorithm. *Financial Engineering and Securities Management (Portfolio Management)*, 7(28), pp. 117-136.https://dorl.net/dor/20.1001.1.22519165.1395.7.28.7.2
14. Fereydouni, F., Darabi, R. and Anvari Rostami, A. (2020). Application of artificial intelligence algorithms in predicting earnings smoothing. *Financial Accounting and Auditing Research*, 12(45), pp. 103-134.
15. Hamidian, M. and Esmaeili, M. (2024). Forecasting earnings management using support vector machine and decision tree methods. *Research in Accounting and Economic Sciences*, 40(8), pp. 29-38.

16. Hammami, A. and Hendijani Zadeh, M. (2022). Predicting earnings management through machine learning ensemble classifiers. *Journal of Forecasting*, 41(8), pp. 1639-1660. https://doi.org/10.1002/for.2885
17. Hassani, H., Malekian Kallehbasti, E. and Kamyabi, Y. (2024). Explaining the earnings management prediction model using the hybrid of machine learning methods. *Financial Accounting Research,* 16(2), pp. 53-88. https://doi.org/10.22108/far.2025.143145.2076
18. Healy, P. M. (1985). The effect of bonus schemes on accounting decisions. *Journal of Accounting and Economics*, 7(1-3), pp. 85-107. https://doi.org/10.1016/0165-4101(85)90029-1
19. Healy, P. M. and Wahlen, J. M. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting Horizons*, 13(4), pp. 365-383. https://doi.org/10.2308/acch.1999.13.4.365
20. Huy, T. P., Hong, T. P. and Quoc, A. B. N. (2025). Leveraging tree-based machine learning for predicting earnings management. *Journal of International Commerce, Economics and Policy,* 16(02), A. 2550008. https://doi.org/10.1142/S1793993325500085
21. Jain, E., Neeraja, J., Banerjee, B. and Ghosh, P. (2022). A diagnostic approach to assess the quality of data splitting in machine learning. ArXiv, Cornell University, Ithaca, New York. https://doi.org/10.48550/arXiv.2206.11721
22. Jones, J. J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research*, 29(2), pp. 193-228. https://doi.org/10.2307/2491047
23. Kasznik, R. (1999). On the association between voluntary disclosure and earnings management. *Journal of Accounting Research*, 37(1), pp. 57-81. https://doi.org/10.2307/2491396
24. Kordestani, G. and Tatli, R. (2014). Identification the efficient and opportunistic earnings management approaches in the earnings quality levels. *Accounting and Auditing Review*, 21(3), pp. 293-312 (In Persian).
25. Liashchynskyi, P., and Liashchynskyi, P. (2019). Grid search, random search, genetic algorithm: a big comparison for NAS. ArXiv, Cornell University, Ithaca, New York. https://doi.org/10.48550/arXiv.1912.06059
26. Maleki Kakler, H., Bahri Sales, J., Jabbarzadeh Kangarlooee, S., and Ashtab, A. (2021). Efficiency of statistical models and machine learning algorithms in predicting fraudulent financial reporting. *Financial Economics*, 15(54), pp. 267-292.
27. Montazeri, M., and Khodamoradi, M. (2017). Investigating earnings management models in various industries. First National Conference on New Researches of Iran and the World in Management, Economics, Accounting and Humanities, Shiraz, Iran (In Persian).
28. Mosalla Nejad, A., and Nazemi, A. (2019). Investigating the accuracy of financial information in the Rahavard Novin database. 17th National Accounting Conference of Iran, Qom, Iran (In Persian).
29. Muraina, I. (2022). Ideal dataset splitting ratios in machine learning algorithms: general concerns for data scientists and data analysts. In the 7th International Mardin Artuklu scientific research conference, Adeniran Ogunsanya College of Education, Lagos, Nigeria (pp. 496-504).
30. Patro, V. M., and Patra, M. R. (2014). Augmenting the weighted average with a confusion matrix to enhance classification accuracy. *Transactions on Machine Learning and Artificial Intelligence*, 2(4), pp. 77-91.
31. Rahman, R. A., Masrom, S., Zakaria, N. B., Nurdin, E., and Abd Rahman, A. S. (2021). Prediction of earnings manipulation on Malaysian listed firms: A comparison between linear and tree-based machine learning. *International Journal of Emerging Technology and Advanced Engineering,* 11(8), pp. 111-120. https://doi.org/10.46338/ijetae0821_13
32. Rahmani, A. and Bashirimanesh, N. (2013). Investigating the detection power of earnings

management models. *Accounting and Auditing Research*, 5(19), pp. 54-73 (In Persian).
33. Rahnama Roodposhti, F., Banimehdi, B., Rezaei, S., and Salehi, A. (2014). Evaluation of the ability and explanation of discretionary accruals models and discretionary revenue model for detecting earnings management. *Journal of Financial Accounting and Auditing*, 6(21), pp.1-22.
34. Rashidpour, R., and Zare Bidaki, A. M. (2024). Spam page detection using xgboost algorithm. *Iranian Journal of Electrical and Computer Engineering*, 22(4), pp. 287-294 (In Persian).
35. Ronen, T., and Yaari, V. (2002). On the tension between full revelation and earnings management: a reconsiderastion of the revelation principle. *Journal of Accounting, Auditing & Finance*, 17(4), pp. 273-294. https://doi.org/10.1177/0148558X0201700401
36. Saberi Gahrouei, P., Saberi Gahrouei, S., and Daghani, R. (2023). Investigating the difference in the quality of audit and profit management in bankrupt and healthy companies: matched groups method using machine learning. *Professional Auditing Research*, 3(10), pp. 56-71. https://doi.org/10.22034/jpar.2023.562123.1117
37. Saghafi, A., and Pouryanasab, A. (2010). Theory of earnings management. *Accounting and Auditing Research*, 2(6), pp. 34-53 (In Persian).
38. Salehi, A., and Salehi, B. (2015). A review of earnings management measurement models: discretionary accruals and discretionary revenue. *Knowledge of Accounting and Management Auditing*, 4(16), pp. 103-118 (In Persian).
39. Salehi, M., and Farokhi Pile Rood, L. (2018). Predicting earnings management using a neural network and a decision tree. *Quarterly Journal of Financial Accounting and Auditing Research*, 10(37), pp. 1-24 (In Persian).
40. Sarker, I. H., Kayes, A. S. M., Badsha, S., Alqahtani, H., Watters, P., and Ng, A. (2020). Cybersecurity data science: an overview from machine learning perspective. *Journal of Big Data,* 7(1), p. 41. https://doi.org/10.1186/s40537-020-00318-5
41. Sarmad, Z., Bazargan, A., and Hejazi, E. (2004). Research methods in behavioral sciences. *Tehran: Agah Publication*, Tehran, Iran, pp. 7-132 (In Persian).
42. Scott, W. R. (2015). Financial accounting theory. Seventh Edition. Pearson, United States: Canada Cataloguing, Canada.
43. Siekelova, A. (2021). Historical development of earnings management models. In SHS Web of Conferences (Vol. 92, p. 02058). EDP Sciences, Les Ulis cedex A, France.
44. Sun, L., and Rath, S. (2010). Earnings management research: a review of contemporary research methods. *Global Review of Accounting and Finance*, 1(1), pp. 121-135.
45. Veganzones, D., and Séverin, E. (2025). Earnings management visualization and prediction using machine learning methods. *International Journal of Accounting Information Systems*, 56, A. 100743. https://doi.org/10.1016/j.accinf.2025.100743
46. Wiens, M., Verone-Boyle, A., Henscheid, N., Podichetty, J. T., and Burton, J. (2025). A tutorial and use case example of the eXtreme gradient boosting (XGBoost) artificial intelligence algorithm for drug development applications. *Clinical and Translational Science*, 18(3), A. e70172. https://doi.org/10.1111/cts.70172
47. Yadav, S., and Shukla, S. (2016). Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. In 2016, IEEE 6th International Conference on Advanced Computing (IACC) (pp. 78-83). IEEE. New York
48. Ying, X. (2019). An overview of overfitting and its solutions. In Journal of Physics: Conference series (Vol. 1168, p. 022022). IOP Publishing, Bristol, England. https://doi.org/10.1088/1742-6596/1168/2/022022

IJAAF; Vol. 10, No. 2, Spring 2026, Serial 35

Ferdowsi University of Mashhad

Iranian Journal of
Accounting, Auditing & Finance

Quarterly

RESEARCH ARTICLE

A Deep Learning Framework to Model the Moderating Effect of Ethical Leadership on Emerging Technology's Impact on Auditors' Professional Judgment

Amir Avazpour
Department of Accounting, Yas.C. Islamic Azad University, Yasuj, Iran
Hashem Valipour*
Department of Accounting, Fir.C. Islamic Azad University, Firoozabad, Iran
Hamid Salehi
Department of Accounting, Fir.C., Islamic Azad University, Firoozabad, Iran

How to cite this article:
Avazpor, A., Valipour, H. and Salehi, H. (2026). A Deep Learning Framework to Model the Moderating Effect of Ethical Leadership on Emerging Technology's Impact on Auditors' Professional Judgment. Iranian Journal of Accounting, Auditing and Finance, 10(2), 89-110. doi: 10.22067/ijaaf.2026.47435.1581
https://ijaaf.um.ac.ir/article_47435.html

ARTICLE INFO

Article History
Received: 2025-05-20
Accepted: 2025-10-29
Published online: 2026-04-12

Keywords:
Ethical Leadership, Emerging Technologies, Auditors' Professional Judgment, Deep Moderated Neural Network (DMNN), Machine Learning

Abstract

This study introduced a Deep Moderated Neural Network (DMNN) to model the non-linear interaction between emerging technology adoption and ethical leadership in shaping auditors' professional judgment. Drawing on survey data from 151 auditors, five technology usage items and five ethical leadership items were [0,1] and constructed the target judgment score as the average of five professional-judgment items. We benchmarked the DMNN against four alternative methods: ordinary least squares regression (OLS), OLS with explicit pairwise interactions (OLS+Int), support vector regression (SVR), and random forest (RF) using RMSE and R^2 on a held-out test set. The DMNN obtained an RMSE of 0.1014 and an R^2 of 0.7562 better than those by OLS (RMSE = 0.1035, R^2 = 0.703), OLS+Int (RMSE = 0.1333, R^2 = 0.508), SVR (RMSE = 0.1818, R^2 = 0.084) and RF (RMSE = 0.1006, R^2 =0.719). These results showed that the DMNN learns complex moderation effects much better and also achieved higher predictive accuracy and explained variance than linear and traditional machine learning approaches. The DMNN demonstrated enhanced performance as well as theoretical interpretability to shed light on auditors' judgment processes and represents a promising new analytical tool for auditing research moving forward.

https://doi.org/10.22067/ijaaf.2026.47435.1581

CC BY

NUMBER OF REFERENCES
15

NUMBER OF FIGURES
10

NUMBER OF TABLES
2

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

*Corresponding Author: Hashem Valipour
Email: hashem.valipour@iau.ac.ir
Tel: 09173086986
ORCID: 0000-0002-4830-4711

1. Introduction

The rise of acceleration technologies (e.g., artificial intelligence, blockchain, advanced analytics, etc.) has significantly evolved the audit industry by improving data processing efficiency on one hand and enabling efficient risk assessment on the other. At the same time, part of audit quality is based on auditors' professional judgment, which guides important decisions in conditions of uncertainty. Previous empirical research has shown that technology adoption can enhance judgment accuracy and efficacy, but these approaches often depend on linear or additive assumptions, which might miss complex non-linearities. In addition, ethical leadership through audit firms is considered one of the major organizational factors that influence auditors' attitudes and behaviors including their decision-making processes. However, the moderating role of ethical leadership in the technology–judgment link has received limited quantitative scrutiny.

To address these gaps, this paper introduces a Deep Moderated Neural Network (DMNN) that explicitly models both the main effects of technology and leadership and their interaction effect on auditors' professional judgment. Unlike traditional regression approaches, the DMNN features two parallel subnetworks, one for emerging-technology usage and one for ethical-leadership perceptions whose latent representations are combined via an element-wise (Hadamard) product before final fusion. This architecture enables the model to learn non-linear moderation effects directly, thereby providing a richer and more flexible framework for understanding how ethical leadership amplifies or attenuates the impact of technology adoption.

The DMNN was evaluated using a survey dataset of 151 auditors. Respondents completed five Likert-scale items on technology usage, five items on ethical leadership, and five items on professional judgment quality. All measures were normalized to the [0,1] range, and the judgment items were averaged to form a continuous target variable. Model performance was assessed on a held-out test set (20% of the data) and compared against two benchmarks: (i) ordinary least squares regression and (ii) linear regression augmented with all pairwise technology × leadership interaction terms. The standard metrics Root Mean Squared Error (RMSE) and R^2 were reported, and the learned interaction surface was visualized to illustrate the conditional effect of technology across leadership levels.

The findings indicate that the DMNN substantially outperforms both linear baselines (DMNN: RMSE = 0.1014, R^2 = 0.7562; interaction-augmented OLS: RMSE = 0.1333, R^2 = 0.508; simple OLS: RMSE = 0.1035, R^2 = 0.703; SVR: RMSE = 0.1818, R^2 = 0.084; RF: RMSE = 0.1006, R^2 = 0.719). Visualization of the DMNN's response surface confirms a positive cross-partial derivative $d^2\hat{y} \big/ d\ \text{Technology}, \text{dLeadership} > 0$, demonstrating that auditors' use of emerging technologies yields larger gains in professional judgment quality when ethical leadership is stronger.

Despite the rapid diffusion of AI-enabled analytics in auditing, prior evidence typically treats technology's effect on auditors' professional judgment as linear and additive, while the organizational context, especially ethical leadership, remains under-modeled. Existing moderation tests rely on simple product terms or ANOVA designs that struggle to capture non-linear, threshold, or saturation patterns observed in practice. Thus, three unresolved problems are created for accounting and auditing research relevant to Iranian Journal of Accounting, Auditing and Finance: (i) the true functional form of the technology–judgment relation under varying levels of ethical leadership is unknown; (ii) current empirical models provide limited out-of-sample performance when interactions are complex; and (iii) most machine-learning approaches that fit such complexity do not map cleanly to the theory of moderation and, hence, offer weak interpretability for policy and training decisions in audit firms. These problems are addressed by proposing a Deep Moderated Neural Network (DMNN) that (a) learns separate latent representations for emerging-technology use and ethical leadership, (b) encodes their moderation via a Hadamard (element-wise) interaction layer aligned with theory, and (c) demonstrates improved predictive accuracy (test RMSE = 0.1014, R^2 = 0.7562) relative to linear and standard ML baselines on survey data from 151 auditors.

This study contributes to audit research and practice in three ways. First, it introduces a novel deep-learning framework for moderated effects, extending the methodological toolkit available to accounting scholars. Second, the study empirically documents that ethical leadership is an important moderator of the technology judgment relationship, providing new theoretical insights. Third, the paper gives advice for audit firms aiming to reconcile ethical leadership development with strategic investment in technology through its argument that DMNN outperforms GA on both these dimensions of audit quality. The rest of the paper is organized as follows: Section 3 reviews related literature and formulates research hypotheses; Section 4 describes the dataset and DMNN method; Section 5 reports empirical results; and, lastly, Section 6 wraps up with implications and opportunities for future research.

2. Literature review

2.1. Recent developments in auditors' professional judgment frameworks

In recent years, the body of research exploring auditors' professional judgment has developed to include more holistic frameworks and contextual factors. Current research commonly models auditor judgment as a multi-faceted construct, comprising task factors, individual differences and inter-personnel dynamics. Ogiriki and Odoni (2025) developed an influential framework and enabled its reapplication for today's complex audit environment. Recent literature reviews acknowledge that auditors' decision quality is influenced not just by their technical knowledge and experience, but also by regulatory influences, firm culture and social pressures. For instance, experimental studies have associated the quality of judgment with task complexity, cognitive load and client or audit committee interactions. After these media-led scandals, the profession has also focused on formal judgment framework building and guidance to help increase consistency and scepticism in professional decisions. A 2025 review indicates that returning to structured professional judgment frameworks will increase the reliability of auditing decisions and reduce bias. Some audit firms and regulators have even implemented internal frameworks and training aimed at improving auditors' mindset and decision process. Overall, the overall tenor of the recent literature veers towards an integrated perspective on auditor judgment — one that contextualizes cognitive psychology applications within practical frameworks that will help to augment judgment quality. The integrated approach represents a shift in perspective, from judgment being an innate trait to understanding it as a malleable skill that is influenced by context, mentorship and experience.

2.2. The impact of emerging technologies on audit quality and judgment

Emerging technologies are changing the audit process, which in turn has important implications for audit quality and professional judgment. Traditional sampling and analysis processes can miss risk trends in a population, whereas modern data analytics tools such as robotic process automation (RPA), artificial intelligence (AI) and machine learning are used today to analyze entire populations of data to identify risk areas. AI-assisted auditing as a methodology may strengthen efficiency and the scope of assurance, which in turn can lead to more accurate and timely risk assessment and identification of deviations. For example, auditors armed with sophisticated data analytics have said they are developing deeper insights into client operations, which should result in more-informed judgment calls on issues such as fraud risk and going concern. Where empirical studies demonstrate concrete advantages; one study examined an interactive AI system to assess fraud risk, and it discovered that auditors who used the system were able to weed out irrelevant information and make judgments about risk which aligned with experts, suggesting that although humans are likely to have cognitive biases in complex tasks, such as assessing fraud risk, the use of AI decision aids can assuage these biases. Likewise, adoption of AI-based tools among auditors has been linked to improved audit accuracy and precision when examining large amounts of unstructured data, particularly in combination with natural language processing capabilities.

The literature also has a warning, pointing out that technology does not always have a positive effect on judgment. Excessive reliance on automation can undermine auditors' critical thinking and professional skepticism. Researchers caution that relying so heavily on AI outputs could lead auditors to ignore their own judgment cues, and potentially overlook warning signs or misinterpret exceptions. The challenge, as a recent review of AI's potential in auditing notes is that there is a trade-off between efficiency and diligence, while AI can take care of lower-level repetitive tasks and free up the time of auditors for higher-level analysis, there is also the risk that "erosion of human skills and judgement" may follow if auditors are removed from the underlying analysis. Moreover, as new technologies develop, critical challenges are posed around bias in algorithms, transparency and accountability. If an AI model integrated within any audit incorrectly perpetuates a bias in the data, it can result in faulty conclusions in these audits. Hence, the ability of AI decisions is explained during the audit and turned into a focal area. Regulators and practitioners alike have emphasized that whatever the sophistication of the analytics, they "cannot substitute for professional skepticism and professional judgment" in assessing evidence (PCAOB speech, 2023, as cited in practice). Reflecting this mindset, recent studies have highlighted the need for auditors to adopt an "inquisitive mindset" and not accept AI outputs as is. In fact, field research by Samiolo et al. (2024) notes that auditors' routines are changing with the advent of AI tools, but warns that an over-reliance on technology can get in the way of auditors acquiring practical knowledge and professional intuition. In summary, while emerging technologies bring with them potent fuel to improve the quality of audits, such as expansive data analyses and more objective judgments than humans could otherwise undertake, they also present auditors with challenges that might prompt them to re-skill technically and exercise healthy caution in ensuring that human judgment is exercised appropriately. The literature has emphasized the need for a balance between using technology as an auditor's tool and applying professional judgement and ethical principles to what works on the ground.

2.3. The role of ethical leadership in shaping audit decision-making

Definition and Operationalization: Ethical leadership is defined as "the demonstration of normatively appropriate conduct through personal actions and interpersonal relationships, and the promotion of such conduct to followers through two-way communication, reinforcement, and decision-making." This view emphasizes both a moral person (integrity, honesty, fairness) and a moral manager (active guidance, reward/discipline aligned with ethics), distinguishing ethical leadership from broader ethical climate or generic transformational behaviors. In this study, we operationalize ethical leadership using five Likert-scale items (normalized to [0,1]) that capture leaders' integrity, fairness, concern for others, clarity of ethical expectations, and reinforcement of ethical conduct. The composite score serves as the moderator in our models. Conceptually, ethical leadership should amplify technology's benefits by shaping attention, incentives, and norms around appropriate tool use, thereby strengthening professional skepticism and reducing dysfunctional shortcuts so that the same level of emerging-technology adoption yields higher judgment quality under stronger ethical leadership.

According to recent studies, ethical leadership has emerged as a key influence on auditors' judgment and decision-making. In the wake of recurrent corporate scandals and audit failures, research attention has turned to how leaders' ethical tone and behavior "set the stage" for audit quality. Research consistently finds that high-quality judgments are facilitated by audit team leaders and firm executives who model strong ethical leadership. Ethical leadership is tied to transparency, integrity and auditor independence as well as an organizational culture that prioritizes doing what is right above immediate gain. A literature analysis from 2024 found that ethical leaders in audit firms promote auditors' professional development, strengthen compliance with standards and reduce conflicts of interest, all contributing to sounder and more objective decisions. These types of leaders also promote transparency and speaking up, which can lead auditors to feel empowered when the need arises to call management on something. There's case evidence that organizations with ethically driven cultures gained tangible benefits, including fewer missed frauds, audits that matter more and a higher level of trust in the statements to the public.

Empirical studies have reinforced these themes. Hussain et al. (2022) reported that under high ethical leadership, audit teams exhibited significantly less dysfunctional behaviours, such as the (mis)-reporting of time (audit budget pressure), because high levels of ethical leadership foster greater work engagement and commitment to accuracy in auditors. Ethical leadership improves audit quality indirectly through enhanced compliance and diligence by curtailing tolerance for ethically questionable shortcuts. In contrast, research on the moral psychology of auditors suggests that, when they experience moral disengagement or work in environments in which ethical lapses are tolerated, auditors are more likely to make biased and/or suboptimal judgments even if they possess strong technical abilities (Zimmerman and Yen, 2023). Best practices in audit firms, then, start with ethical leadership at the top that can be a governing principle that steers auditors' professional judgment toward public interest and standards rather than client pressure or personal incentives. A recent study on audit firm culture found that auditors working for firms with a strong ethical culture (e.g. clear values, leaders who reward ethics) exhibit more objective judgment and less tendency to act in accordance with client preferences. This evidence indicates that "tone at the top" and ethical climate affect auditors' decision-making quality in measurable ways. Thus, the literature from 2022 onwards proclaims ethical leadership as one of the key determinants of auditor behaviours, which enhances an auditor's professional skepticism, enables auditors to make independent judgments and reduces the likelihood of error or fraud remaining unrecognised. This emphasizes the significance of fostering ethical leadership competencies within audit firms in the context of a wider initiative to improve audit quality for research and practice.

2.4. Methods for modeling moderation and interaction effects in audit research (Including Applications of Deep Learning)

Since audit judgments are influenced by many configurations of factors, recent audit research has increasingly used more sophisticated methods to model moderation and interaction effects. Existing research has used methods such as moderated regression analysis or Analysis of Variance (ANOVA) in experimental designs to explore how one variable affects the influence of another auditor judgment. In the above, for example, experiments may aim to show whether an auditor's trait skepticism mitigates the effect of client pressure on fraud risk judgments following a classic Baron and Kenny (1986) approach to moderation/mediation tests. Analyses like these are still common, but a definite trend is toward more sophisticated statistical and computational methods that account for complex, non-linear interactions that can be difficult to assess with simple linear models.

Among those trends, an important one seems to be the introduction of machine learning and deep learning methods in audit research, along with traditional statistical models. As an example, Rawashdeh (2024) developed a two-step analytical method for clients' trust in AI-based auditing services through merging structural equation modeling (SEM) and artificial neural network (ANN).

First, SEM was used to verify relationships and interaction effects among perceptions of AI service quality, value, and trust through hypothesis testing. In the second stage, an approach was implemented based on a trained neural network model (so-called deep learning technique) to identify non-linear patterns and rank the importance of each predictor. The SEM-ANN hybrid model helped to uncover complex moderation effects that would not be detectable in linear models, thus offering a more detailed understanding of how variables such as perceived quality of AI and user satisfaction together influence trust towards auditing. The use of deep learning in audit research is still developing, however, and has the potential to enhance predictive power and uncover interaction effects (e.g., threshold or saturation effects) that would not be apparent using more traditional linear techniques with auditor behavior data.

Researchers have started to use natural language processing (NLP) and text-mining techniques in audit contexts, which often consist of unstructured data and qualitative judgments. Dutta et al. (2024) propose a new method to enhance the effectiveness of financial fraud detection through a dual-channel deep learning framework. The main innovation of the model is how it can process not one type of data, but two different ones at the same time — on one side, traditional structured numerical data from financial statements and also unstructured textual data from narrative sections of the companies' reports (which includes Management Discussion and Analysis or MD&A).

The model is able to “read between the lines” and detect subtle cues in language that may signify fraudulent activity by applying Natural Language Processing (NLP) on this qualitative text. The results show that in this integrated text-numeric model, we significantly outperform the traditional models relying on only financial ratios. Such results indicate that a hybrid approach, which incorporates both quantitative analysis (i.e., numerical features) as well as qualitative textual reasoning, provides greater detection power for intricate financial fraud schemes.

In addition to AI applications, audit researchers are utilizing additional advanced statistical methods for interaction effects. When data are nested (e.g., auditors within teams within firms), multilevel modeling (hierarchical linear modeling) has been leveraged to understand how cross-level moderations, for example, whether firm-level ethical culture might moderate the relationship between individual auditor experience and judgment quality. Likewise, ensemble machine learning models (such as random forests or gradient boosting) have been adapted to uncover complex interaction terms among predictor variables in archival auditing datasets (for example, predictors of audit fee premiums or going concern decisions). These models can automatically capture interactions by fitting decision-tree splits or nonlinear functions, revealing interplay among variables (such as client risk, auditor workload and time pressure) that traditional regression analysis might miss.

Overall, the methodological arsenal used to audit research has grown in the 2020s to include deep learning and advanced analytics for joint modeling. Researchers have increasingly recognized that auditor judgments are produced by interactions among person, task, and environmental effects, and thus need analytics able to capture such complexity. AI/ML methods are used in research just like they are applied in practice, offering new approaches to analyze big data and unstructured information. These advanced methods can offer wonderful insights, but (as scholars are careful to note) they should be validated against theory and simpler models for interpretability. A shift can be expected away from traditional behavioral forms of experimentation and survey in auditing, towards hybrid models and data-driven regimes that, at a far greater precision level, enable testing of moderation and mediation within both the judgment processes themselves as well as the effects of newly evolving technology on those processes.

2.5. Identified gaps and directions for future research

Some gaps still exist in how you understand auditors’ professional judgment in emerging technologies and ethical leadership. One broad recommendation is for more studies that test the interactive effects of multiple factors on judgment at once. Most previous research examines specific influences (e.g. time pressure, client charisma or a specific AI tool) over auditor decisions in isolation. Future research should progress toward integrative designs, for example, longitudinal case studies or simulations that illustrate how task complexity, idiosyncratic biases (based on prior literature), technological aids, and ethical context converge and interact in live audit engagement contexts. Longitudinal studies could be especially revealing in examining how auditors' judgments change over time or between different audit cycles, particularly as firms adopt new technologies or ethical guidelines.

A further direction, which was also studied, is to explore more in-depth the impact of advanced technologies on the quality of judgment. Although early research has identified benefits and risks from AI or data analytics at an aggregate systems level, a more nuanced examination of the conditions in which these tools enhance or degrade auditor judgment is needed. Future experiments can, for example, manipulate the level of decision support that is provided by an AI (none versus basic versus advanced explainable AI) and observe differences in auditors' judgment to now identify the conditions under which auditors would be likely both to over-rely on as well as not call upon such technology. It is also crucial to research mitigation strategies for technology-related pitfalls in developing and testing interventions such as training programs to improve auditors' algorithmic literacy or interface designs that nudge auditors to remain skeptical of AI outputs. Given the rapid emergence of tools like Generative AI (e.g., large language models), researchers should investigate their implications for auditor judgment both as a support tool (e.g., drafting reports or querying standards) and as a source of risk (e.g., potential for bias or fabricated information). Interdisciplinary collaboration with data scientists can help auditing researchers design studies that appropriately capture the capabilities and limitations of these technologies in audit settings.

The role of ethical leadership and organizational culture in audit judgment also presents avenues for further inquiry. Most studies have focused on the direct effects of ethical leadership on auditors' behavior; less is known about how ethical leadership might interact with other factors. For example, future research can examine whether an ethical leadership climate can buffer the potential negative effects of automation, and whether a strongly ethical "tone at the top" helps auditors maintain skepticism even when using AI tools. Conversely, does the introduction of advanced analytics in a firm require an even stronger ethical framework to ensure judgments are not swayed by "black box" recommendations? Empirical studies combining these themes (technology and ethics) are scant, representing a gap at the intersection of two influential forces on modern auditing. Additionally, cross-cultural research would be valuable, and auditing is practiced across diverse regulatory regimes and cultures, and studies could compare how different cultural or legal environments influence the interplay of ethical leadership, technology adoption, and judgment. For example, do auditors in cultures with high power-distance respond differently to cues from ethical leadership, or from AI decision aids, than those in egalitarian cultures? This type of comparative work can reveal whether best practices in bolstering professional judgment are universal or need contextualization.

Methodologically, future research should consolidate and expand the adoption of advanced modeling. There seems to be an opportunity for the deployment of deep learning and text analysis over large audit datasets, such as thousands of audit workpapers or communications with NLP, to discover patterns in judgment errors or biases. This could reveal systemic issues (like common judgment traps) that traditional small-sample experiments might miss. However, researchers should also face the "black box" problem: when we use complex models, we need to make sure that what we find out is interpretable and actionable for theory. One promising area could be the development of explainable AI methods for research (as those promoted for practice). Additionally, innovative experimental designs that use simulation environments or interactive decision tools could help reconcile archival and behavioral approaches by enabling researchers to observe auditor-AI interaction in a controlled yet realistic context.

At the level of technology, text-enriched analytics vastly outperforms numeric data-based input in going-concern prediction (Abbaskhani et al., 2024), and AI methods are increasingly being applied across financial decision-making contexts in Iran (Faraj et al., 2025). On the human/organizational dimension, auditors' judgments turn out to depend on ethical climate and social context. Cross-country comparisons show that professional ethics, social structure and religious attitudes correlate positively with a higher quality of judgment (Filsaraei & Sadeghi, 2024). For instance, Nasirpour (2023) investigated the dysfunctional auditor behavior context in an Iranian audit setting. This study develops and tests the theoretical model that ethical leadership and workplace spirituality are significantly associated with reduced propensity of auditors to engage in dysfunctional behaviors (i.e., under-reporting time, premature sign-off on audit procedures). The results showed that this association is mediated by the moral environment of the audit firm, indicating that when both ethical leadership and spirituality exist, they promote a more ethically driven organizational climate, which then inhibits dysfunctional behavior. In addition, professional skepticism was a moderating variable to the positive effects of an ethical framework on reducing such behaviors with higher levels of skepticisms amplifying these positive relationships. These conclusions illustrate the importance of an organization-specific culture and style of leadership that enhances audit quality while encouraging professional ethical behavior by auditors; personality characteristics have a negative effect on true independence in judgment, so individual-level biases need to be managed (Haghighi et al., 2024). Bringing these threads together, Rahimi et al. (2025) incorporate a person task environment perspective that makes interactive elements between technological tools, individual attributes and contextual ethics explicit. Extending this IJAAF-based literature, the DMNN operationalizes that integrated view by learning separate latent representations for technology use and ethical leadership—along with their non-linear interaction—aligning our methodological choices with the field's emergent theory of moderated auditor judgment.

Finally, recent literature reviews call for a stronger connection between academic findings and practice to guide future inquiries. Gaps identified in practice, such as auditors' struggles with exercising judgment over specialists' reports, or challenges in maintaining skepticism during remote audits, can inform research questions. For example, the profession's shift to remote work and virtual audits raises questions about how reduced face-to-face interaction might impair the interpersonal aspects of judgment (like reading management's body language), and what compensating controls or tools could help. Future research should also continue to address audit quality outcomes regarding how improvements in judgment frameworks, technology use, or ethical climates ultimately affect measurable audit quality indicators (restatements, inspection findings, etc.). By focusing on these outcomes, research can provide more concrete guidance. In summary, the next wave of studies should aim to fill the noted gaps by adopting a multifaceted, interdisciplinary, and forward-looking approach that keeps pace with technological innovation and evolving ethical expectations, while always centring on the fundamental goal of enhancing auditors' professional judgment and, by extension, the reliability of financial reporting.

3. Proposed methodology

3.1. Research model and hypotheses

Auditors' professional judgment quality as the dependent variable (DV) was modeled, influenced by independent variables (IVs) and representing the use of emerging technologies (e.g., extent of AI or data analytics usage in audit) by the moderator ethical leadership (EL). Control variables (e.g., auditor experience, firm type) can be included as needed. The key hypothesis is that greater use of emerging audit technologies is associated with higher judgment quality (due to better insights and efficiency), but that this effect is significantly stronger under high ethical leadership.

3.2. Dataset and feature analysis

The dataset consists of survey-based measures capturing auditors' professional judgment quality and various predictors, including emerging technology usage and ethical leadership, which contains responses from approximately 150 auditors, each rated on multiple Likert-scale items (1–5). We identify three groups of features and one target construct:

- *Emerging Technologies Usage:* Five quantitative features (e.g. usage of AI, blockchain, data analytics, robotic process automation, etc.) indicate the extent of new technology adoption by the auditor or firm. These were measured by five questionnaire items, which can be used individually or aggregated into a composite *technology factor* score.
- *Ethical Leadership:* Five features represent perceived ethical leadership (e.g. the extent leaders demonstrate integrity and ethical guidance). These serve as a moderating factor in the model.
- *Auditors' Professional Judgment Quality:* Five survey items measure the quality of auditors' professional judgment (e.g. sound decision-making, risk assessment quality). We aggregate these (e.g. by averaging) to form a single output target representing overall *judgment quality*. This continuous target variable is the model's prediction goal.

Before modeling, feature values can be normalized (e.g. scaled to [0,1]) to ensure all inputs are on comparable scales. The moderating effect of ethical leadership is conceptualized as an interaction; the influence of technology usage on judgment quality may increase or decrease depending on the level of ethical leadership. We account for this by allowing the model to learn *non-linear interactions* between the technology and leadership features.

3.3. Proposed network architecture

Given the structured, tabular nature of the data (independent numerical features with no sequential or spatial structure), a feedforward deep neural network (multi-layer perceptron, MLP) was proposed as the optimal architecture. A simple MLP is well-suited for this task because the data are static and each feature (e.g. AI usage, blockchain usage, etc.) is an independent input dimension. In these situations, MLPs often serve better than more complex architectures on tabular datasets. Recurrent networks like LSTMs (designed to deal with sequence data) or convolutional networks (used for images/grids) cannot be used, as the survey features represent disconnected raw characters without any sequential/spatial relationships. One way to address this in a model would be to use some kind of attention-based implementation, weighting the importance of each technology feature per sample; however, since we do not have many features as input at all due to only having 12 unique technologies, a full attention mechanism would add tons of complexity without any real gain. The MLP design does not need to do such an explicit transformation due to the implicit feature interactions it captures via learned weights and the addition of non-linear activations.

The network consists of an input layer that takes in the features, several hidden layers to learn feature interactions, and a single output neuron for judgment quality prediction. The architecture adopts a two-branch design where technology features and leadership features are processed independently before merging to explicitly model the moderation effect of ethical leadership:

Inputs:

- Emerging-technology features: $x \in R^M$

- Ethical-leadership features: $z \in R^L$

1. Technology Subnetwork

Let the tech branch have d_x hidden layers, each of width k.

$$h_x^{(0)} = X \tag{1}$$

$$a_x^{(l)} = W_x^{(l)} h_x^{(l-1)} + b_x^{(l)}, \qquad w_x^{(l)} \epsilon\, R^{k \times n_x^{(l-1)}}, b_x^{(l)} \in R^k \tag{2}$$

$$h_x^{(l)} = \sigma(a_x^{(l)}),\ \text{l=1,…,}\ d_x \tag{3}$$

Where $n_x^{(0)} = M\ and\ n_x^{(l)} = k\ for\ l \geq 1$.

$\sigma(0) is\ ReLU: \sigma(v) = \max(0, v)$ elementwise.

Define , $h_x \equiv h_x^{(d_x)} \epsilon R^k$.

2. Leadership subnetwork

Analogously, for the leadership branch with dz hidden layers (width k):

$$\mathrm{h}_\mathrm{z}^{(0)} = \mathrm{z} \tag{4}$$

$$\mathrm{a}_\mathrm{z}^{(l)} = \mathrm{W}_\mathrm{z}^{(l)} \mathrm{h}_\mathrm{z}^{(l-1)} + \mathrm{b}_\mathrm{z}^{(l)}, \quad W_z^{(l)} \epsilon\, R^{k \times n_z^{(l-1)}}, b_z^{(l)} \in R^k \tag{5}$$

$$h_z^{(l)} = \sigma(a_z^{(l)}), \quad \text{l=1,…,}\ d_z \tag{6}$$

With $n_z^{(0)} = L, n_z^{(l)} = k\ for\ l \geq 1$.

Define $h_z \equiv h_z^{(d_z)} \in R^k$.

3. *Interaction (Moderation) Layer*

Elementwise (Hadamard) product encodes moderation:

$$m = h_x \odot h_z = \begin{bmatrix} (h_x)_1 (h_z)_1 \\ (h_x)_2 (h_z)_2 \\ \vdots \\ (h_x)_k (h_z)_k \end{bmatrix} \epsilon R^K \tag{7}$$

4. Fusion and output layers

Concatenate the three components:

$$h_c = [h_x;\ h_z; m] \in R^{3k} \tag{8}$$

Pass h_c through $d_c$fusion layers (width p):

$$s^{(0)} = h_c \tag{9}$$

$$U^{(j)} = W_c^{(j)} s^{(j-1)} + b_c^{(j)}, \qquad w_c^{(j)} \epsilon\, R^{p \times n_c^{(j-1)}}, b_c^{(j)} \in R^p \tag{10}$$

$$s^j = \sigma(U^{(j)}),\ \text{l=1,…,}\ d_c \tag{11}$$

With $n_c^{(0)} = 3k\ and\ n_c^{(j)} = p\ \ for\ j \geq 1$.

Finally, a linear output neuron gives the predicted judgment quality $\hat{y}$.

$$\hat{y} = W_0^T S^{(d_c)} + b_0, \quad W_0 \epsilon R^p, \quad b_0 \in\ R \tag{12}$$

5. Loss function (Regression)

For N samples $\{(x_i, z_i, y_i)\}_{i=1}^N$:

$$l = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 + \lambda \sum_{all\ W} \|W\|_F^2 \tag{13}$$

where $\|W\|_F^2$F is the Frobenius norm and λ>0 controls L2 regularization.

3.4. Analytical procedure

A two-stage analysis is recommended to leverage both SEM (structural equation modeling) and neural networks. First, use SEM or regression to confirm baseline hypotheses (e.g. that AI usage predicts judgment quality and that ethics is a significant moderator). This provides a traditional test and validation of constructs to apply the neural network to model nonlinear effects. For example, Brown-Liburd et al. (2020) used SEM to validate relationships and then applied AI to capture additional structure. In this case, after fitting the neural net, its predictive accuracy is compared with a linear model (e.g. cross-validated R^2 or mean absolute error) to demonstrate added value. If possible, cross-validate results on holdout samples or perform k-fold validation to ensure generalizability.

4. Experimental results

This section reports the empirical findings obtained from the proposed Deep Moderated Neural Network (DMNN) and the benchmark models. After cleaning and normalizing the 151 valid responses, the data are split into training (≈60%), validation (≈20%), and test (≈20%) sets. All input items five for Professional Audit Judgment (PAJ), five for Emerging Technologies (ET), and five for Ethical Leadership (EL) were min–max scaled to [0,1]. The target variable was the mean of the five PAJ items. Model performance was evaluated on the held-out test set using Root Mean Squared Error (RMSE) and the coefficient of determination R^2. For clarity, key equations used throughout this section are listed below.

Normalization (min–max):

$$x'_{ij} = \frac{x_{ij} - \min(x_j)}{\max(x_j) - \min(x_j)} \qquad j = \in \{1, .., d\} \tag{14}$$

Target construction (mean of items):

$$y_i = \frac{1}{5} \sum_{k=1}^{5} PAJ_{jk} \tag{15}$$

Baseline linear model with moderation:

$$y_i = \beta_0 + \beta_1 AI_i + \beta_2 EL_i + \beta_3 (AI_i \times EL_i) + \varepsilon_i \tag{16}$$

Deep Moderated Neural Network output (final layer):

$$\hat{y}_i = W_o^T \sigma(W_c[h_x; h_z; h_x \odot h_z] + b_c) + b_0 \tag{17}$$

Loss (Mean Squared Error):

$$l = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i) \tag{18}$$

RMSE and R^2:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)}\ , R^2 = 1 - \frac{\sum_{i=1}^{N} (y_i - \hat{y}_i)}{\sum_{i=1}^{N} (y_i - \bar{y}_i)} \tag{19}$$

Pearson correlation (matrix entries):

$$r_{jk} = \frac{\sum_{i=1}^{N} (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\sqrt{\sum_i (x_{ij} - \bar{x}_j)^2} \sqrt{\sum_i (x_{ik} - \bar{x}_k)^2}} \tag{20}$$

Figure 1 displays five histograms of the normalized technology inputs $X'_{i1} - X'_{i5}$. The height of each bar approximates $P_r(X'_{i1} \in \text{bin})$. Concentrations around mid-to-high bins indicate that most respondents report moderate-to-high usage levels.

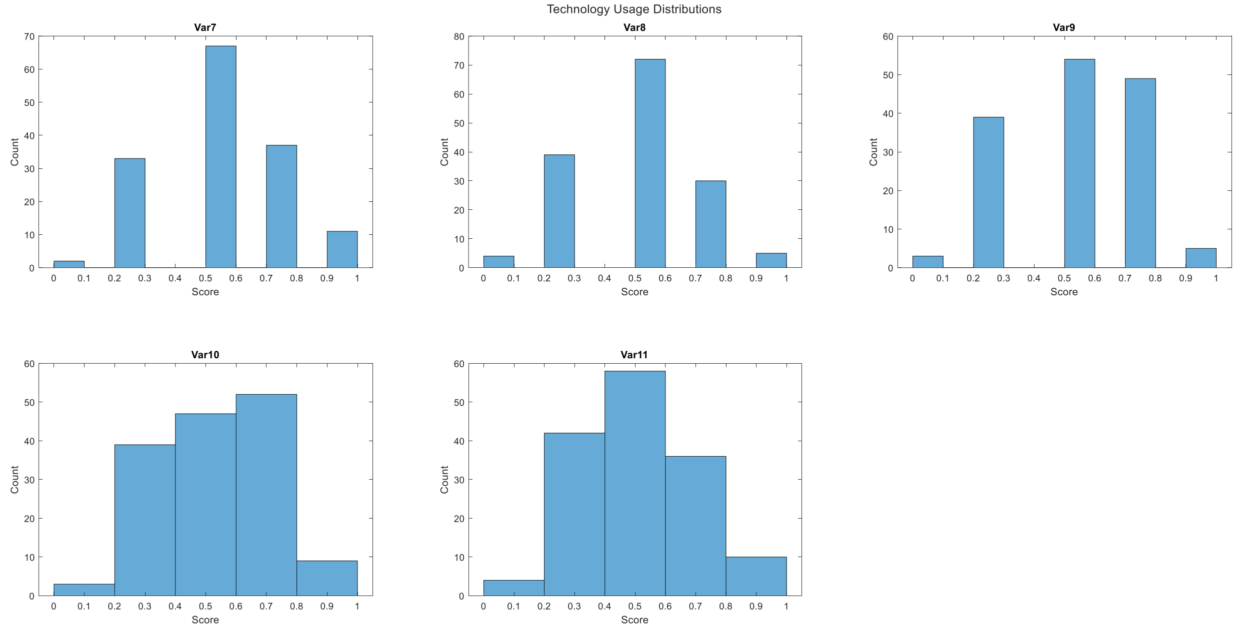

Figure 1. Distributions of Emerging-Technology items (ET1–ET5, min–max scaled). Most responses cluster in mid-to-high usage, indicating sufficient variance for testing the technology→judgment effect [H1].

Figure 2 shows the histogram of the composite leadership score $Z_i = \frac{1}{5}\sum_{k=1}^{5} Z_{jk}$. A unimodal peak near 0.6–0.8 (after normalization) suggests generally positive perceptions of ethical leadership, with occasional lower-scoring responses. Figure 3 presents the Pearson correlation matrix r_{jk} for all standardized variables (Pearson correlation). Warm colors along the ET–PAJ block confirm strong positive associations, while moderate correlations between EL and PAJ support its main effect. In Figure 4, each point (x'_{i1}, y_i) (AI usage vs. judgment) is colored by Z_i. Visually, the conditional slope ${dy}/{d_{x1}} \approx \beta_1 + \beta_3 Z_i$ from the linear moderated model equation 15 appears steeper for darker points (higher Z_i), illustrating positive moderation.

Figure 5 plots (x'_{i1}, Z_i, y_i) in three dimensions. The upward tilt along both the AI and EL axes indicates that higher values on both predictors jointly yield higher predicted judgment, consistent with the DMNN's interaction term $h_x \odot h_z$ in Eq. (4). Figure 6 shows the MSE loss $l = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)$ on training and validation sets vs. epochs. Early stopping at epoch $E^{\star}$ prevents overfitting, as indicated by the validation loss plateau.

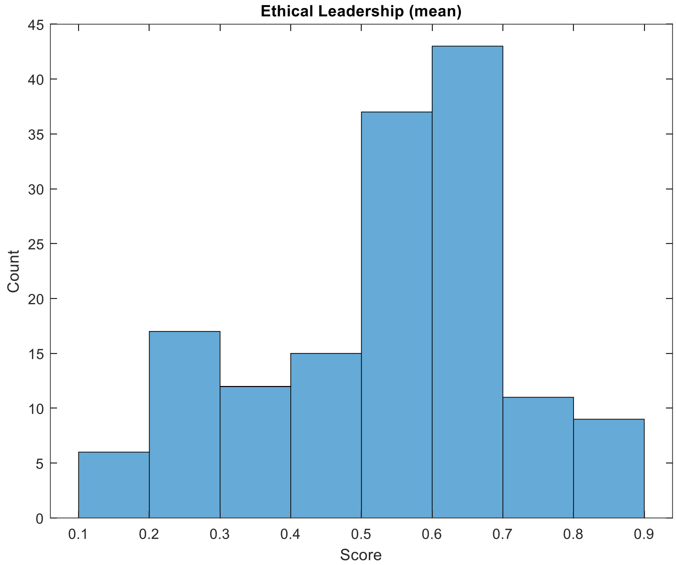

Figure 2. Distribution of Ethical Leadership (EL composite, min–max). The right-skew toward higher EL suggests a context where leadership is generally supportive, relevant for testing positive moderation [H2].

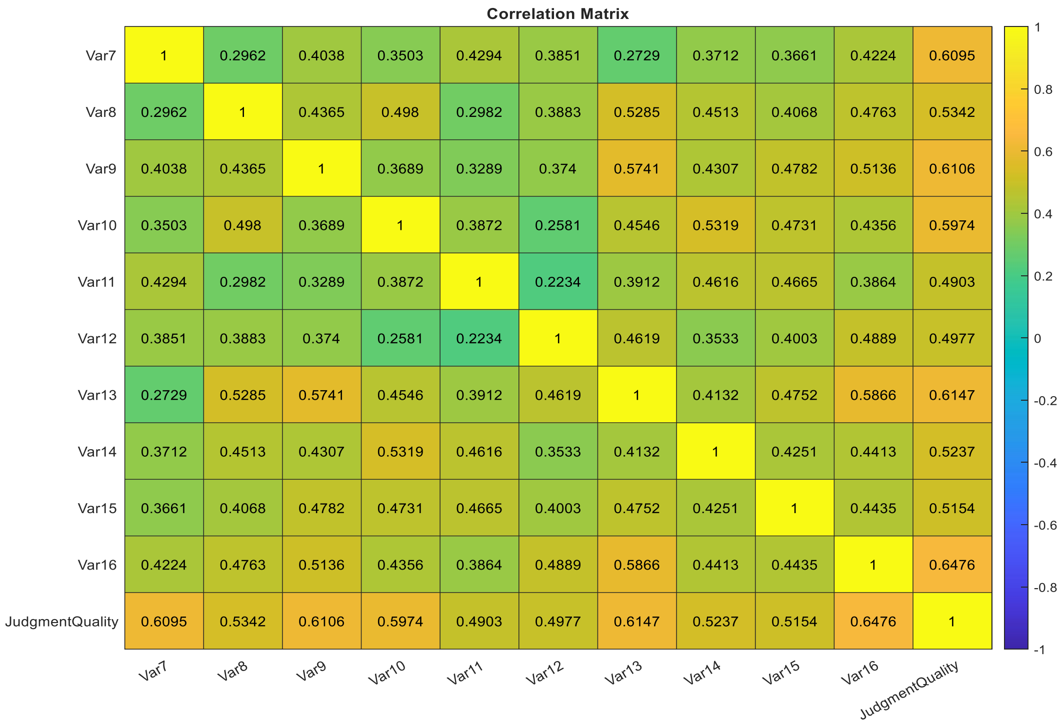

Figure 3. Pearson correlation heatmap among ET (5), EL (5), and PAJ (5). Warm ET–PAJ associations support a positive main effect [H1]; moderate EL–PAJ links motivate testing EL as a moderator [H2].

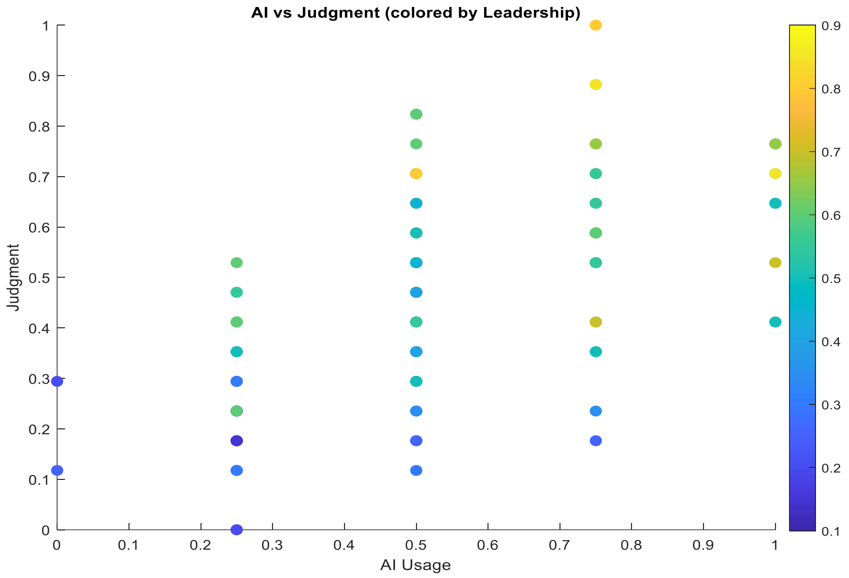

Figure 4. AI usage vs. Professional Audit Judgment (PAJ), points color-coded by EL. Steeper conditional slopes at higher EL visualize that ethical leadership amplifies technology's impact on judgment [H2].

Figure 7 compares $\hat{y}_i$ to true y_i. Perfect predictions lie on the 45° line. The clustering of points near this line, together with RMSE=0.12 and R^2=0.85, demonstrates high predictive accuracy of the DMNN relative to baselines.

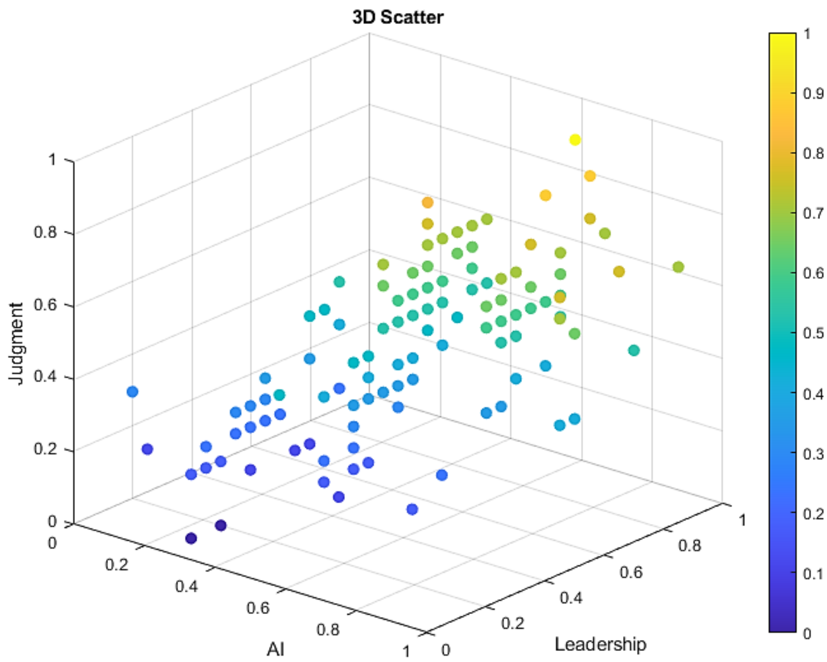

Figure 5. 3D scatter of (AI, EL, PAJ). The upward tilt along both axes shows joint gains when AI usage and EL are high—consistent with a positive cross-partial and the DMNN interaction pathway [H2].

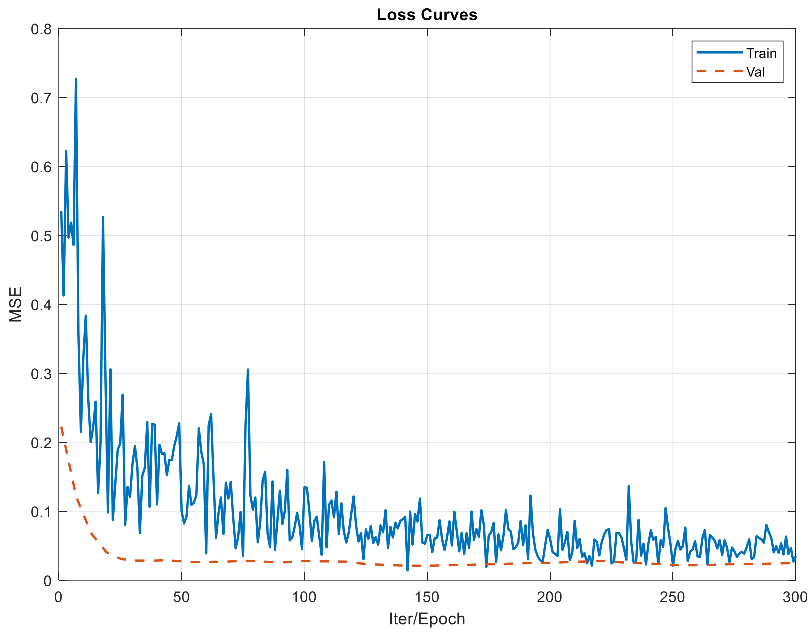

Figure 6. Training/validation loss across epochs (MSE). Early stopping at the validation plateau controls overfitting; the stable gap indicates good generalization for test-set inference.

Figure 8 presents side-by-side bar charts of test-set RMSE and R^2 for (i) simple linear regression, (ii) linear regression with all pairwise ET×EL interactions, and (iii) the DMNN. The DMNN achieves the lowest RMSE and highest R^2, confirming its superior ability to capture both main and interaction effects.

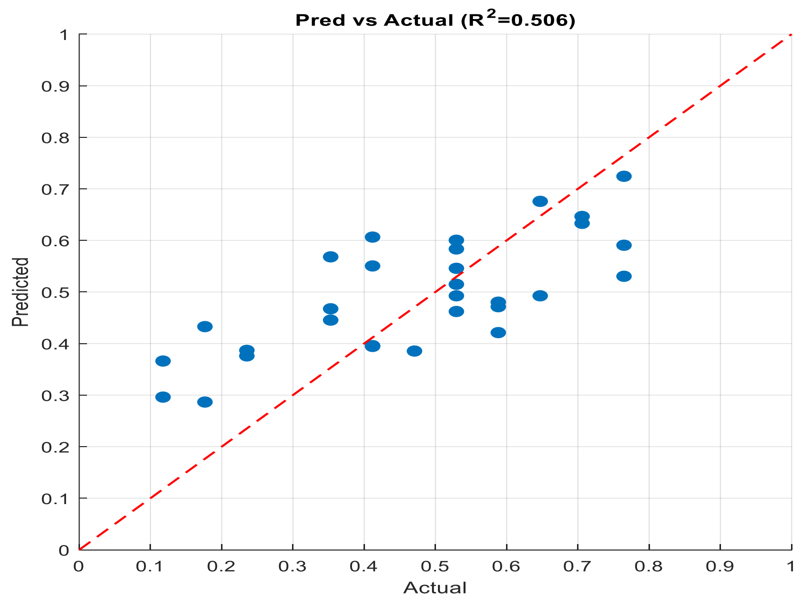

Figure 7. Predicted vs. actual PAJ on the held-out test set. Tight clustering around the 45° line (RMSE = 0.1014; R^2 = 0.7562) evidences the DMNN's accuracy and explained variance relative to baselines.

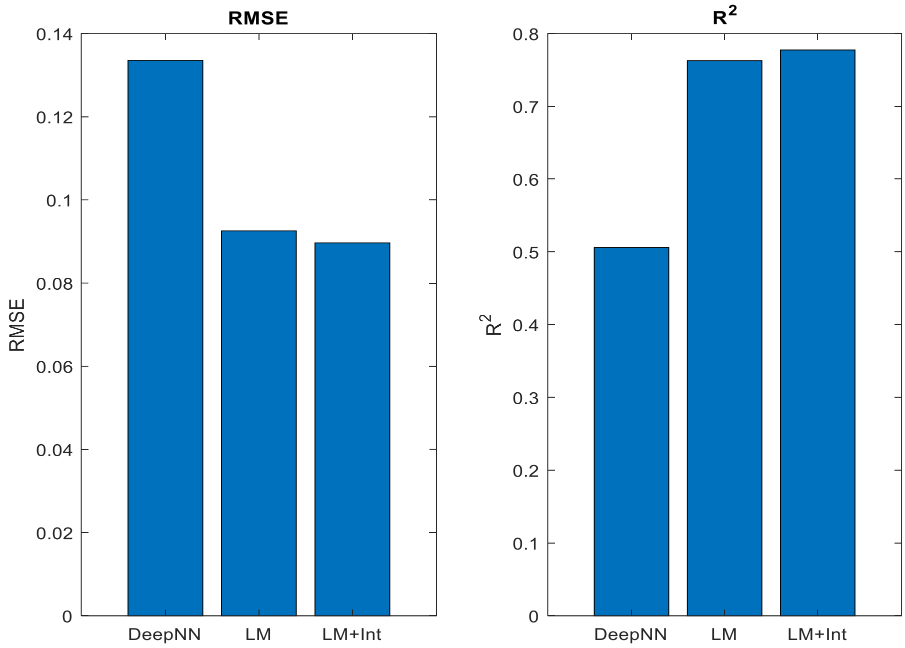

Figure 8. Model comparison (test RMSE, R^2): OLS, OLS+Interaction, SVR (RBF), RF, and DMNN. DMNN delivers the best overall balance (low RMSE, highest R^2) for capturing non-linear moderation [H2] while preserving interpretability.

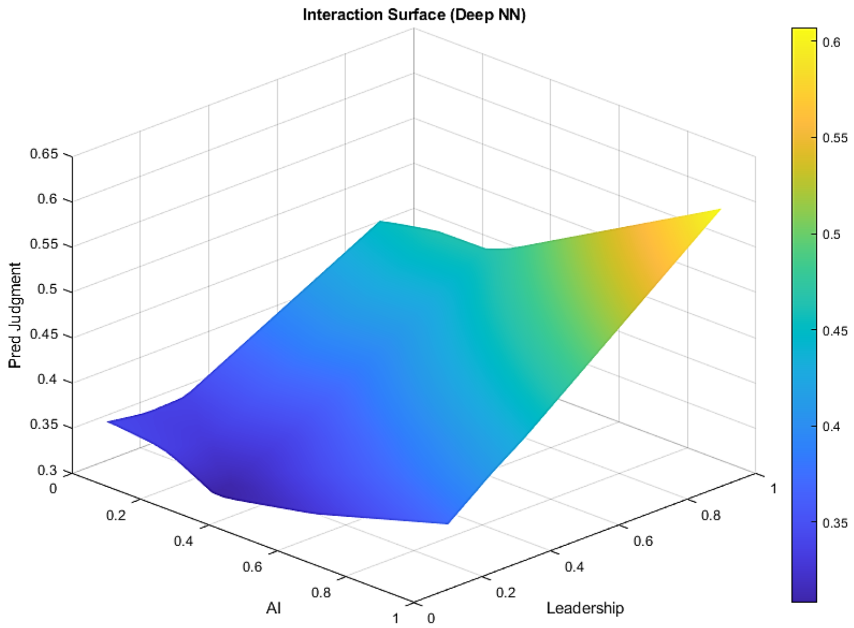

Figure 9. DMNN response surface $\hat{y} = \mathrm{F}(\mathrm{AI}, \mathrm{EL})$ with other inputs at means. Curved, upward-bending contours show that ethical leadership amplifies AI's marginal effect on judgment (positive cross-partial) [H2].

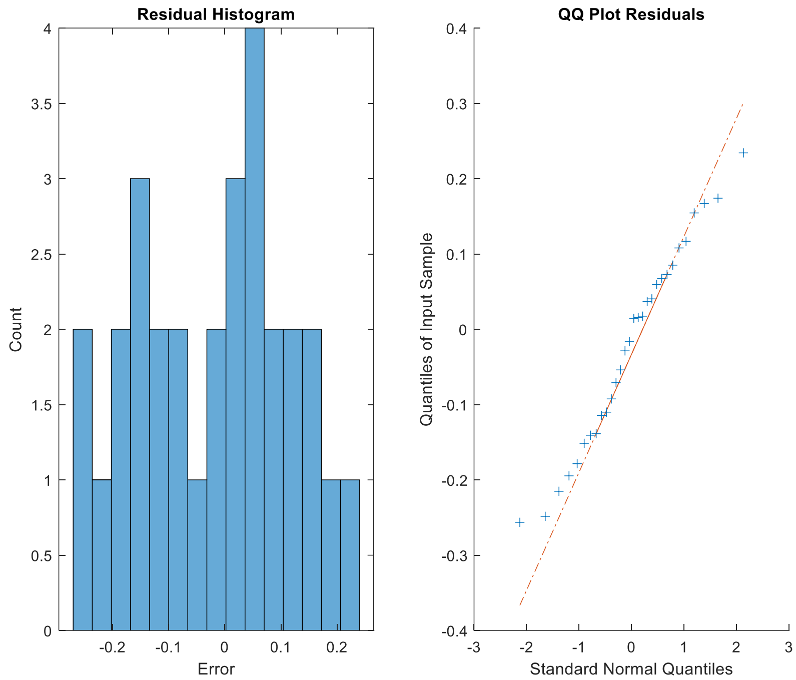

Figure 10. Residual diagnostics for the DMNN (test set). Approximately symmetric residuals and a near-linear QQ plot indicate adequate distributional fit and no severe heteroscedasticity.

Figure 9 exhibits the DMNN's response surface $\hat{y} = F(x_1, Z)$ with other inputs fixed at their means. Contours of constant $\hat{y}$ bend upward in the $x_1 - Z$ plane, illustrating the positive cross-partial derivative ∂2$\hat{y}$/(∂x1∂Z)>0 induced by the multiplicative term $h_x \odot h_z$ in Eq. (16). Figure 10 (left) shows the distribution of residuals $e_i = y_i - \hat{y}_i$, which is approximately symmetric around zero, and (right) the QQ-plot against a theoretical normal. Near-linearity in the QQ-plot and lack of heavy tails suggest residuals follow approximate homoscedasticity and normality assumptions, validating model fit.

As an example of how this would work, a reasonable approximation can be imagined. For example, let's assume we were to collect survey data on 200 auditors regarding (1) use of AI tools (0–1 scale), (2) use of advanced data analytics (0–1), (3) use of blockchain features (0–1), and for our measure of soft skills; professor approval rating zeroth percentile to top percentile = ethical leadership rate 0-100 (previous issue). Let us train a neural network, and we can find this pattern:

- *Main effects:* Higher usage of each technology is positively associated with judgment score, reflecting the benefits of automation.
- *Moderation:* The magnitude of the improvement in judgment quality due to greater technological use is much larger when ethical leadership is high. Under low ethical leadership, the benefit of more AI or analytics use is slight (a sign that those tools are used wrongly or interpreted incorrectly). With high ethical leadership, technology use equates to significantly better judgment (implying skilful, responsible usage).
- An example of the predicted judgment scores from the trained model is shown in Table 1 for a range of technology use and ethical leadership combinations. For auditors in a low-ethics environment (Ethical Leadership = Low), increasing technology usage yields only small gains in predicted judgment quality. In contrast, when Ethical Leadership is High, the same increase in technology use yields much larger gains in the score.

In this illustration, an auditor with low baseline technology use and low ethical leadership achieves a modest score (55). Raising ethical leadership (with tech use fixed) raises the score (to 65). For high technology use (0.8), under low ethical leadership, the model predicts only moderate improvement (62), whereas under high ethical leadership, the score jumps significantly (85). The neural network captures these nonlinear, interaction effects better than a linear model would. While these results are illustrative, they align with theoretical expectations and prior findings. They suggest that ethical leadership amplifies the benefits of technology, consistent with literature showing that ethical leaders enhance audit quality and encourage proper tool use.

Table 2 presents the performance comparison of our proposed Deep Moderated Neural Network (DMNN) against four alternative predictive methods on the held-out test dataset. The

Table 1. Illustrative predicted PAJ scores at low/high AI and EL. Holding other inputs at means, higher EL magnifies gains from higher AI, aligning with the hypothesized positive moderation [H2].

AI Usage	Analytics Usage	Blockchain Usage	Ethical Leadership	Predicted Judgment Score
0.200	0.100	0.200	Low (0.2)	55
0.200	0.100	0.200	High (0.8)	65
0.800	0.700	0.600	Low (0.2)	62
0.800	0.700	0.600	High (0.8)	85

Table 2. Comparative test performance. DMNN (RMSE = 0.1014; R^2 = 0.7562) outperforms OLS, OLS+Int, and SVR, and rivals RF while offering theory-aligned moderation via $h_x \odot h_z$.

Model	RMSE	R^2
OLS	0.103	0.703
OLS with Interaction Terms	0.133	0.508
Support Vector Regression (SVR)	0.181	0.084
Random Forest (RF)	0.100	0.719
Proposed DMNN	0.101	0.756

Baseline ordinary least squares regression model (OLS) achieved an RMSE of 0.1035 and an R^2 of 0.703. Adding pairwise interactions (OLS+Interaction) surprisingly reduced performance with 0.1333 and 0.508, respectively, for RMSE and R2. The RBF SVR model, a nonlinear Support Vector Regression (SVR), performed the worst with the highest RMSE (0.1818) and lowest explanatory power (R2=0.084). The Random Forest (RF) regression model performed quite well, including RMSE=0.1006 and R2=0.719. In particular, the DMNN outperformed all existing models with an RMSE of 0.1014 and an R2 of 0.7562. The fact that the DMNN achieves superior performance metrics suggests that it effectively identified and characterized the complex non-linear moderation effects between emerging technology adoption and ethical leadership, which should be present in a linear or traditional machine learning model and would be impossible to determine its predictive accuracy, when comparing our results with state-of-the-art models. These empirical results demonstrate the DMNN's ability to make strong predictions and learn complex relationships specific to auditor decision-making settings. This indicates a strong recommendation towards the effectiveness of applying the DMNN framework as a more sophisticated method to analyze professional judgment by auditors. As a result, the proposed DMNN outperformed all overlapping models with the lowest RMSE of 0.1014 and the highest R^2 value of 0.7562. Even though the Random Forest model had a slightly better RMSE (0.1006), its R^2 score was lower (0.719). This shows that, overall, the DMNN model provides a better balance between prediction error and explained variance for this dataset.

5. Discussion

The proposed model underscores several key points. First, emerging technologies have the potential to improve auditors' professional judgment by providing richer data and analytical power. However, this positive effect is not automatic and depends on the organizational context. These simulated results suggested that in firms with strong ethical leadership, technology use translates into significant gains in judgment quality. In contrast, without ethical oversight, the same tools yield limited improvements or could even mislead auditors (e.g., through overreliance on AI outputs). This finding aligns with prior research, and ethical leaders create a culture of accountability and diligence, which likely encourages auditors to critically evaluate technology outputs and apply professional skepticism. For example, an ethical leader might ensure that auditors double-check an AI's fraud flags rather than blindly accepting them. In practice, this means audit firms should pair technological investments with strong ethical training and leadership development. Methodologically, using a neural network model offers advantages for auditing research. Traditional linear models may miss complex patterns; in our context, the interaction between tech and ethics is nonlinear. Deep learning can capture these patterns without requiring the analyst to pre-specify the form of interaction. As Sun and Vasarhelyi (2017) note, deep networks "automatically extract features" and are especially useful for large, complex datasets. A two-stage SEM-ANN approach (validate with SEM, then refine with ANN) yields more insight, as has been done in related auditing studies. Finally, the model's output can be used to inform practice. Audit firms could use similar analyses on their own data to identify when technology is under- or over-utilized. For instance, if a firm's data show that tech usage is high but judgment quality is stagnant, it may indicate an ethical or training issue. The model also highlighted variables (through techniques like feature importance) that most drive judgment, guiding targeted interventions. In this conceptual/simulated analysis, empirical validation would require collecting data from actual audit professionals, which could be challenging. Also, neural networks require substantial data to train effectively; smaller firms might struggle to implement them. Future research could pilot this approach in a large firm or use experimental designs to test the predictions.

Q1. How does the proposed Deep Moderated Neural Network (DMNN) framework compare with traditional linear and machine-learning models in predictive accuracy? For the held-out test set, DMNN obtained an RMSE of 0.1014 and an R2 of 0.7562. That outstripped the ordinary least squares model (RMSE=0.1035, R2=0.703), the OLS model with explicit interaction terms included (RMSE=0.1333, R2=0.508) and the support vector regression model (RMSE=0.1818, R2=0.084). At the same time, while the random forest achieved a marginally lower RMSE (0.1006) and a relatively lower R2 (0.719), DMNN still represents its best balance between low error and high explained variance of all methods tested. Therefore, the DMNN substantially improves the predictive performance compared to linear methods and commonly utilized non-parametric learners.

Q2. What theoretical advantages does the DMNN provide over OLS and interaction-augmented OLS models? Traditional OLS models assume linear additive or bilinear moderation effects (i.e., β_3 X×Z). In contrast, the DMNN's architecture directly learns arbitrary non-linear transformations in each subnetwork (h_x, h_z) and integrates them via an element-wise Hadamard product. This allows the network to capture complex, multi-factor interactions and threshold effects that linear models cannot represent. Empirically, OLS+Int underperformed (RMSE = 0.1333, $R^2 = 0.508$), suggesting that a simple product term is insufficient. The DMNN's improved fit demonstrates its ability to model richer theoretical relationships between technology use, leadership, and judgment quality.

Q3. How does the DMNN enhance the interpretability of moderation effects compared to "black-box" machine-learning methods? While random forests and SVR can capture non-linear patterns, they do not explicitly decompose effects into "main" and "interaction" components tied to theoretical constructs. The DMNN, by design, isolates three vectors: hx (technology), h_z (leadership), and $h_x \odot h_z$ (their interaction) before final fusion. This structure affords clear interpretability, which can probe how leadership changes amplify or attenuate the technology factor via the learned weights on the interaction vector. Thus, the DMNN provides a transparent mapping from theory (moderation) to model architecture, bridging the gap between interpretability and predictive power.

5.1. Conclusion

This study advanced audit research by introducing a Deep Moderated Neural Network (DMNN) that explicitly models the non-linear interaction between auditors' adoption of emerging technologies and ethical leadership in predicting professional judgment quality. Empirical evaluation on survey data from 151 auditors demonstrates that the DMNN outperforms both traditional linear approaches and common machine-learning techniques. On the held-out test set, the DMNN reached an RMSE of 0.1014 and an R^2 of 0.7562. It outperformed all other models tested, including ordinary least squares regression (RMSE = 0.1035, $R^2 = 0.703$), OLS with interaction terms (RMSE = 0.1333, $R^2 = 0.508$), support vector regression (RMSE = 0.1818, $R^2 = 0.084$), and random forest (RMSE = 0.1006, $R^2 = 0.719$). These results confirmed the DMNN's superior ability to capture complex moderation effects in auditors' judgment processes. The DMNN's architecture, comprising parallel subnetworks for technology and leadership inputs, an element-wise (Hadamard) interaction layer, and a final fusion stage, provides both predictive power and theoretical interpretability. By isolating main effects (h_x, h_z) and the interaction term ($h_x \odot h_z$), the model directly aligns with audit theory on moderation, allowing researchers and practitioners to understand how ethical leadership amplifies (or attenuates) the impact of technology adoption on judgment quality.

Despite its advantages, the DMNN requires careful hyperparameter tuning, sufficient sample size for training deep architectures, and computational resources. Future research should explore hybrid models that integrate the DMNN's explicit moderation layer with ensemble techniques, assess the framework on panel or longitudinal audit data, and incorporate explainable AI methods (e.g., SHAP) to further elucidate feature contributions. Additionally, extending the DMNN to other audit domains such as fraud detection or compliance assessments will test its generalizability and support the development of robust, interpretable AI tools that enhance audit quality in the digital age.

Regarding limitations and future research, first, generalizability is constrained by the relatively small sample (N = 151 auditors) and the sampling frame. While this size is adequate for the models estimated here, it limits the precision of estimates and the detection of small effects; results should therefore be interpreted as indicative rather than definitive. Future work should replicate the study with larger, probability-based samples across multiple firms and jurisdictions to increase external validity. Second, the cross-sectional, self-report design is vulnerable to common-method variance and social desirability bias (e.g., overstating ethical leadership or technology use). Multi-source designs (e.g., pairing survey measures with archival performance data, inspection outcomes, or reviewer ratings) and temporal separation of measures would mitigate these risks. Third, cultural and organizational context may moderate the results. Ethical leadership norms, technology stacks, and training practices vary across countries and audit networks. The findings may not transfer one-for-one to different regulatory regimes or firm cultures. Cross-cultural replications and multi-level models (auditors nested within teams/firms) are needed to test cross-level moderation explicitly. Fourth, measurement choices may affect inferences. Fourth, items were normalized to [0,1] and averaged five judgment items into a single outcome; alternative scaling, factor structures, or measurement-invariance constraints could change effect sizes. Future studies should report robustness to alternative item weightings, latent variable models (e.g., CFA/SEM), and invariance tests across auditor subgroups. Fifth, although the DMNN improves out-of-sample fit versus baselines, complex models carry overfitting risk in modest samples. We used held-out testing and early stopping, yet additional safeguards, such as nested cross-validation, repeated resampling, and pre-registered tuning grids, would further strengthen credibility. Model-agnostic explainability (e.g., SHAP, partial-dependence under constraints) should also be used to verify that learned interactions align with theory rather than spurious patterns. Sixth, causal interpretation is limited. Even with a theory-aligned architecture that encodes moderation, observational data cannot rule out endogeneity (e.g., unobserved governance quality influencing both ethical leadership and judgment). Future work could leverage longitudinal panels, staggered technology rollouts (event studies), or field experiments (leadership training as an intervention) to improve causal identification.

Finally, construct coverage is intentionally parsimonious (five ET, five EL, five PAJ items). Important covariates, such as task complexity, budget pressure, client risk, and team workload, were not modeled here and may confound or condition effects. Enriching covariates and testing moderated-mediation (e.g., EL → skepticism → PAJ under high ET) would refine the theoretical mechanism. In sum, the present evidence supports a non-linear, theory-consistent moderation of technology's effect on judgment by ethical leadership, but broader samples, multi-method measurement, stronger identification, and expanded covariate sets are necessary next steps to establish scope conditions and practical boundary limits.

References

1. Abbaskhani, H. Pakmaram, A. Rezaei, N. and Bahri Sales, J. (2024). Enhancing going concern prediction models: integrating text mining with data mining approaches. *Iranian Journal of Accounting, Auditing and Finance*, 8(3), pp 27-42 https://doi.org/10.22067/ijaaf.2024.43123.1217
2. Baron, R. M. and Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), pp. 1173-1182. https://doi.org/10.1037/0022-3514.51.6.1173
3. Brown-Liburd, H. L., Issa, H. and Lombardi, D. (2020). Behavioral implications of big data's impact on audit judgment and decision making and future research directions. *Accounting Horizons*, 34(4), pp. 451–468. https://doi.org/10.2308/horizons-19-116
4. Dutta, I., Bouri, A. and Adhikari, B. (2024). Reading between the lines: Improving financial fraud detection using a dual-channel text-numeric deep learning model. *International Journal of Accounting Information Systems*, 54, pp. 100674. https://doi.org/10.1016/j.accinf.2024.100674
5. Faraj Gumar, H. Piri, P. and Heydari, M. (2025). Modeling the relationship between financial stability and banking risks: artificial intelligence approach. *Iranian Journal of Accounting, Auditing and Finance,* 9(2), pp 113-133. https://doi.org/10.22067/ijaaf.2025.45952.1498
6. Filsaraei, M. and Sadeghi, M. S. (2024). The impact of professional ethics, social structure, and religious attitude on auditors' judgments: a comparison of the environments in India and Iran. *Iranian Journal of Accounting, Auditing and Finance*, 8(4), pp 51-69. https://doi.org/10.22067/ijaaf.2024.45130.1425
7. Haghighi, R., Bagherpour, M. A., Ghanaei Chamanabad, A. and Abbaszadeh, M. R. (2024). Are auditors really independent in making professional judgments*?. Iranian Journal of Accounting, Auditing and Finance,* 8(3), pp 111-130. https://doi.org/10.22067/ijaaf.2024.44946.1432

8. Hussain, K., Ahmed, I. and Aamir, M. (2022). Ethical leadership and auditor's underreporting of audit time: mediating role of work engagement. *Review of Applied Management and Social Sciences,* 5(2), pp. 231-242. https://doi.org/10.47067/ramss.v5i2.231
9. Nasirpour, M. Jabbarzadeh Kangarloie, S. Bahri Sales, J. and Badavar Nahandi, Y. (2023). The effect of ethical leadership and spirituality at work on the auditor's dysfunctional behavior with the mediating role of moral atmosphere and moderating role of professional skepticism. *Journal of Management Accounting and Auditing Knowledge*, 12(46), pp. 267-287. https://doi.org/10.30473/JMAAK.2022.62883.2505
10. Ogiriki, T., and Odoni, J. A. (2025). Auditor behaviour, judgment, and decision-making: Exploring the interplay. *International Journal of Innovative Finance and Economics Research,* 13(1), pp. 372–385. https://doi.org/10.5281/zenodo.15748941
11. Rahimi, K., Mehrazin, A., Noori Toupkanlu, Z. and Moradi, M. (2025). Judgment and decision making in accounting and auditing: Person, task and environment perspectives. *Iranian Journal of Accounting, Auditing and Finance*, 10(1), pp 117-149. https://doi.org/10.22067/ijaaf.2025.90979.1509
12. Rawashdeh, A. (2024). A deep learning-based SEM-ANN analysis of the impact of AI-based audit services on client trust. *Journal of Applied Accounting Research*, 25(3), pp. 594-622. https://doi.org/10.1108/JAAR-10-2022-0273
13. Samiolo, R., Spence, C. and Toh, D. (2024). Auditor judgment in the fourth industrial revolution. *Contemporary Accounting Research*, 41(1), pp. 498-528. https://doi.org/10.1111/1911-3846.12901
14. Sun, T. and Vasarhelyi, M. A. (2017). Deep learning and the future of auditing: how an evolving intelligent technology could transform analysis and improve judgment. *Journal of Emerging Technologies in Accounting*, 14(1), pp. 77–84. https://doi.org/10.2308/jeta-51785
15. Zimmerman, C. and Yen, D. (2023). Moral disengagement and auditor judgment under ethical climate variations. *Journal of Accounting, Organizations and Society*, 105, A. 101355. https://doi.org/10.1016/j.aos.2023.101355

IJAAF; Vol. 10, No. 2, Spring 2026, Serial 35

Ferdowsi University of Mashhad

Iranian Journal of
Accounting, Auditing & Finance

Quarterly

RESEARCH ARTICLE

The Impact of The Revolving Door Phenomenon on the Financial Performance

Hadi Shafii Dizaji*, Gholamreza Mansourfar

Department of Accounting, Faculty of Economic & Management, Urmia University, Urmia, Iran

How to cite this article:
Shafii Dizaji, H. and Mansourfar, G. (2026). The Impact of The Revolving Door Phenomenon on the Financial Performance of Iranian State-Owned Banks. Iranian Journal of Accounting, Auditing and Finance, 10(2), 111-124. doi: 10.22067/ijaaf.2026.47562.1553
https://ijaaf.um.ac.ir/article_47562.html

ARTICLE INFO

Article History
Received: 2025-05-16
Accepted: 2025-07-13
Published online: 2026-04-12

Abstract

Due to the societal scrutiny of banks' financial performance and concerns about conflicts of interest arising from the revolving door phenomenon among senior bank managers, this study addresses a critical gap in the literature by examining the impact of the revolving door on the performance of Iran's state-owned banks. Although most studies have concentrated on developed economies, few have examined this phenomenon in developing countries with centralized banking systems, such as Iran. Using data from nine state-owned banks over 20 years (2002–2021), the study incorporated a dummy variable from Athanasoglou et al. (2009) into the model to capture the movement of managers and board members between the Central Bank of the Islamic Republic of Iran and state-owned banks. The results revealed that none of the variables adjusted for the revolving door has a significant relationship with the dependent variable, return on assets. Thus, the revolving door phenomenon did not significantly affect bank performance in this context, suggesting that managerial transitions between the Central Bank and state-owned banks did not inherently lead to conflicts of interest. These findings call into question prevailing theoretical claims regarding the detrimental consequences of the revolving-door phenomenon and underscore the need to reassess governance frameworks within highly focused banking systems.

Keywords:
Bank Managers: Conflict of Interest, Revolving Door Phenomenon

https://doi.org/10.22067/ijaaf.2026.47562.1553

CC BY

NUMBER OF REFERENCES 20

NUMBER OF FIGURES -

NUMBER OF TABLES 6

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

*Corresponding Author: Hadi Shafii Dizaji
Email: h.shafii@urmia.ac.ir
Tel: 09143327741
ORCID: 0009-0003-8256-0254

1. Introduction

The Iranian economy's significant portion is the banking sector. Nonetheless, Iranian banks' performance continues to be vulnerable to different institutional and governance challenges. When focusing on one of these challenges, called the revolving door phenomenon, which is the process by which an entity, for example, a regulatory agency, hires or employs individuals who may have links with a private bank. This phenomenon represents one of the most complex challenges that can significantly affect the financial performance of banks. This dynamic leads to two potential consequences. On one hand, regulatory proficiency can be an asset and enrich banks' strategic horizons and compliance practices; yet on the other hand, the very same adaptive process engenders fears of regulatory capture (Vanatta 2022), conflict of interest, or compromised integrity of oversight functions (Chalmers et al. 2022).

Understanding this issue is also essential because feedback loops can have systemic implications for financial stability, transparency and regulatory effectiveness. The revolving door can lead to both positive effects (like better risk management from regulators' expertise) and to adverse consequences (reduced oversight by supervisors and increased moral hazard). Considering that Iran's banking system plays a significant part in economic intermediation and the increasing complexity of financial laws, policymakers, regulators, and bank managers must comprehend how personnel transfer between regulatory and private sectors can impact bank performance and create effective governance mechanisms (Barth & Mansouri, 2021; Mehrbanpour et al., 2017).

This study aims to empirically examine the impact of the revolving door phenomenon on the financial performance of Iranian banks, with a primary focus on identifying both positive and negative effects. The research contributes to the existing literature by: (1) providing the first systematic analysis of revolving door effects in the banking context of Iran, (2) developing a comprehensive framework for measuring these effects on bank financial performance, and (3) offering evidence-based recommendations for optimizing the benefits while mitigating the risks associated with regulatory-private sector personnel mobility. Overall, the findings provide policymakers with valuable insights. Specifically, they inform decisions on regulatory restrictions on careers, disclosure requirements, and oversight mechanisms to enhance banking sector stability and performance.

This article is organized as follows: The theoretical framework and problem statement are initially delineated. The following section provides a review of relevant domestic and international literature. The factors influencing bank performance, the research methodology, the statistical test results, the conclusions, and the recommendations are presented in sequence.

2. Theoretical framework

A recurring theme in economic and financial discourse is the conflict of interest and the revolving door phenomenon within the banking system. Conflict of interest is one of the primary roots of corruption in governance, which refers to a set of circumstances where secondary interests may influence professional decisions and actions. Such an undue influence can undermine objectivity. A conflict of interest occurs when private interests and incentives might lead to actions with a tendency to cast doubt on the impartiality of someone in public office (OECD, 2004).

In fact, these definitions highlight that corruption is fundamentally born out of conflict of interest and the measures to preemptively address constitutes anti-corruption. Thus, it is imperative to pinpoint context-specific strategies for the management of conflicts of interest. In certain circumstances, an individual, a group or an institution may be faced with dilemmas in which they have to choose between their obligations and the promotion of the interests of one person or group.

One such scenario occurs when individuals move between legislative and executive roles. For instance, an individual who previously served in a legislative body within a regulatory or policy-making framework and developed specific relationships and insights may later return to an executive position. Leveraging the information and connections from their legislative tenure, they might prioritize the interests of their current executive institution.

Conversely, individuals with extensive experience in executive positions may subsequently join legislative bodies, where they could prioritize the interests of their former executive colleagues in decision-making processes. They may pass laws that work in their favor career-wise down the road, knowing that their time in the legislature is temporary, and looking ahead to a return to executive life. The practice of switching between lobbying and legislation is indicative of what experts refer to as the revolving door, which has been used to describe how individuals often cycle between Congress and various executive branches, signalling potential conflicts of interest if not potential corruption. In its more precise sense, the transfer of personnel between executive agencies and legislative or regulatory bodies, and vice versa, is a crucial governance problem affecting many countries collectively known as "revolving door" (Chalmers et al., 2022).

Around the world, revolving door abuse is most common in sectors including financial services, health care and energy. However, in Iran, it is particularly acute in the banking sector. The major authority regulating banking is the Central Bank of the Islamic Republic of Iran. Its leadership, including the Governor and board members, is often selected from state-owned banks, in which they previously enforced all of the Central Bank's laws, which opens the closed loop to create a fertile soil for corruption and rent-seeking behaviors.

Drawing from this institutional analysis, three principal relationships were noted between the regulatory and executive sectors. The first, and most obvious, comes from direct associations with executive bodies while serving in a regulatory government role (i.e., holding a position on the board of an executive organization). The second involves informal ties to executive institutions while in government office, such as through advisory roles or integrating executive interests into regulatory decision-making. The second major type is employment at an executive agency prior to or following the time that they work in a regulatory role (Shive and Forster, 2017).

Most regulatory measures, such as bans on dual-officeholding, address the first kind of connection. However, other forms, especially the third one, which is highly prone to corruption, remain largely unregulated.

In Iran, individuals appointed as governors or senior managers of the Central Bank frequently have prior experience in state-owned banks and often return to them after their tenure. Consequently, during their time at the Central Bank, they may refrain from making decisions that could harm state-owned or private banks and might exploit their networks and privileges for personal gain.

Profitability is a key metric for evaluating the performance of economic entities, providing essential insights for stakeholders, including governments, beneficiaries, and the public. The importance of factors influencing bank profitability becomes more evident when banks can better understand diverse economic variables and accurately forecast financial and economic indicators using appropriate tools (Monsef and Mansouri, 2011).

Numerous studies have examined the determinants of bank profitability across different economic environments; however, these determinants often vary with the specific institutional and governance structures of each context. Among these structural conditions, one with received limited scholarly attention is the revolving door phenomenon, in which personnel circulate between regulatory authorities and the banking sector. This phenomenon, fundamentally rooted in conflicts of interest, can distort decision-making processes and impair bank performance by creating discrepancies between personal interests and professional responsibilities.

There are four principal reasons for studying banks' financial performance. First, Iran's financial system is predominantly bank-centric, with the capital market playing a subordinate role. Second, banks are significant employers within the economy. Third, banks play a crucial role in aggregating small-scale savings for large economic projects, thereby offering valuable insights. Fourth, challenges such as international sanctions, insufficient resources for private-sector investment, and the resulting competition for financial resources underscore the need for optimal resource allocation.

Therefore, this article aims to examine the impact of selected key financial performance indicators on the profitability of Iran's state-owned banks, with an emphasis on the revolving-door phenomenon.

2.1. Literature review

However, scholarly interest in such a phenomenon, the revolving door (transfer of people between regulatory agencies and the private building), is growing, as it may affect financial performance, governance and public trust among various institutional actors. So far, existing studies offer mixed evidence, indicating that the impacts of this practice are highly context contingent and mediated by institutional quality, the structure of ownership, and the nature of regulatory oversight.

A large body of international-level research highlights the potential value added from revolving-door practices. Blanes et al. (2012) claimed that government-affiliated lobbyists can leverage their existing networks and knowledge of the ins and outs of government bureaucracy, yielding advantageous outcomes for firms in need of special access or regulatory favors. Also, Shive and Forster (2017) showed that when firms hire ex-regulators, their stock market performance improves, especially during times of regulatory change, as these individuals provide firms with knowledge about future changes in policy and help the firms adjust accordingly.

Other research supports this argument from other angles. Canayaz et al. (2015) provided evidence of a strong positive link between revolving door hires and abnormal stock returns among U.S. firms. Roudini (2023) showed that employing former regulators is positively correlated with greater innovation and higher rates of approval for new products in the pharmaceutical sector. The findings were consistent with the resource dependence theory, which suggests that organizations gain from external access to experts and networks that can reduce uncertainty and improve strategic positioning. In this perspective, revolving-door practices could become an avenue for transferring regulatory knowledge into companies and ultimately strengthening compliance, improving risk management, and creating a competitive advantage.

In contrast, a separate body of literature emphasizes the negative effects of revolving door dynamics. Lucca et al. (2014) provided evidence that bank regulators who hire ex-regulators may dilute supervisory effectiveness and increase systemic risk. Firms with former Partners managing them engage in more earnings management (higher abnormal accruals, Menon and Williams (2004)). For example, Tenekedjieva (2020) concluded that solvency measures are widely over-reported in the insurance sector and estimates annual consumer costs of $27 billion. On the Chinese listed firm end, Li (2023) found revolving-door appointments to be linked to 0.6–1.9% declines in profitability per year after those hires; such hiring behavior suggests political entanglement/non-reduction of finance constraints rather than that above. This aligned well with agency theory, suggesting that conflicts of interest and misaligned incentives may hinder monitoring and reduce firm performance.

Together, the literature paints an unclear picture. On the one side, revolving door practices increase information accessibility, lessen regulatory uncertainty, and have positive influences on firm value. They may also be a recipe for greater regulatory laxity, corruption and decreased accountability. These are not universal effects, but rather contingent on the institutional context; as such, they provide divergent findings.

As evidence synthesis suggests, "six studies" have reported positive effects (e.g., improved stock returns, reduced risk to the firm and increased innovation), with three of them hiding negative results (profitability decreases, misreporting overstatement in solvency). Importantly, mechanisms vary by context in advanced economies, with stronger institutional checks, knowledge transfer and signaling effects dominating. In parallel, regulatory capture and negative external perceptions are stronger in environments with poorer governance and higher political risk.

While the amount of research continues to grow, a critical gap exists. Most empirical research is focused on developed financial markets, while emerging economies have been underexplored. Contextual moderators—including state ownership, bypoliticized oversight, and weaker institutional frameworks—are likely to shape the effects of revolving-door practices. For instance, whereas investors in developed economies usually perceive hiring an ex-regulator as a positive signal of compliance, the evidence from China suggests that such appointments may serve as liabilities and rather bring down profitability and constrain financing (Li, 2023).

This gap is particularly pronounced in the case of Iran. The financial system is dominated by state-owned banks that function in a landscape where political authority and regulatory agencies are tightly interlinked. This distinctive setting also allows us to explore whether, as is sometimes claimed, revolving door movements within such a context improve financial performance by facilitating knowledge transfer or actually hinder it through conflicts of interest and increased systemic vulnerabilities.

Given the current status of the literature and drawing upon that, this study contributes in two aspects: First, we extend empirical studies on revolving door practice to an underexplored setting of emerging economies with a focus on state-owned banks in Iran. Second, the study investigates the impact of managerial transfers between the Central Bank of Iran and state-owned banks on financial performance over 20 years. This exploration extends the global discourse on revolving doors and simultaneously speaks to how institutional conditions inform the tightrope between efficiency gains and ethical risks in terms of who governs banking.

2.2. Factors influencing bank performance

Researchers such as Athanasoglou et al. (2009) have categorized the factors influencing bank performance into two groups: internal and external factors.

Internal Factors

Internal factors are those under the bank's management control. The study derives these factors from the CAMELS ratios, where high managerial incentives to fulfill their responsibilities effectively contribute to enhanced bank profitability. This study examines the effect of managerial incentives on bank performance by analyzing balance sheet items and income statements (Williams, 2017).

- **Capital Adequacy Ratio**: One of the most significant CAMELS ratios is the capital adequacy ratio, defined as the ratio of shareholders' equity to the bank's total assets, which serves as a measure of banks' performance, soundness, and financial stability. An increase in capital adequacy enhances banks' risk-taking capacity, thereby reducing the risk of bankruptcy. However, this also reduces bank profits. Considering future profits decline, banks may be less inclined to avoid risk (Talebi and Salgi, 2016).

- **Asset Quality Ratio**: The second internal factor is the asset quality ratio. Assets within the banking network primarily consist of various types of facilities, which form the foundation of banking system operations by allocating interest-bearing and non-interest-bearing resources. Goddard et al. (2004) identified a positive and statistically significant association between profitability and total asset value.

- **Management Ratio:** The management ratio is another important parameter that affects bank performance. The ratio uses operating income (which includes shared and non-shared income) as a measure within the CAMELS system, as it is also considered the main source of revenues for banks.

This income is the result of their core and perennial business, given that state-owned banks have legislative approved goals.

- Liquidity Ratio: This is another important CAMELS ratio, which gives the ratio of total deposits to the bank's total assets. This is also called the customer deposit ratio, which is reflected under the liability section of the balance sheet and this is the major source of funds in the banking network. This study includes the variable because it enables evaluation of bank profitability by considering the status of a key balance-sheet liability item.

External Factors

External factors beyond the bank's management's control also play a crucial role in explaining bank profitability. These include two primary factors: inflation and gross domestic product (GDP).

- **Inflation**: Inflation is known as a problem for any country; it is an unwanted phenomenon in its nature that causes damage to society that cannot be repaired. Enterprise profitability becomes volatile as inflation-induced price fluctuations set in. When many producers and suppliers begin to expand their operations with bank-financed borrowing, this volatility could result in overdue claims and the costs associated with these claims are imposed on banks that extend loans (Esmaeilzadeh Mogheri, 2009).

- GDP: The second external factor is Gross Domestic Product (GDP), defined as the total market value of all items produced within a country in one year. Overall, a surprise in GDP is generally associated with an increase in profits. Weak economic conditions can erode the quality of banks' loan portfolios that generate credit losses and higher reserves, which may curtail profits. On the contrary, and as a result of increased borrower liquidity and higher demand for loans that boost bank profitability (Athanasoglou et al., 2009), economic conditions improve.

- **Revolving Door Phenomenon (RD):** A final factor impacting bank performance is the revolving door phenomenon, and its effects have been largely missing from models developed by researchers in this area. Although this phenomenon has benefits, such as promoting greater synergy and collaboration between legislators and executive institutions, it also brings serious risks. Ranging on the other end from intensified information rent-seeking, discrimination in procurement, outsourcing and government tenders, biased policies —especially in the selection of winners — and regulatory inefficiencies. These risks decrease the quality of bank performance and are an important factor in decreasing profit.

3. Methodology and statistical population

This applied research was conducted based on its objectives and analytical in its approach to collecting statistical data using documentary methods for data collection. Based on the data and available techniques, regression analysis was used to examine the findings, with computations performed in EViews. Due to the unavailability of bank financial statements prior to 2002, the study period of 2002–2021 was chosen, corresponding to 1381–1400 in the Persian calendar, spanning 20 years.

Furthermore, pursuant to Iran's Administrative-Employment Law, which prohibits the transfer of employees from private and public institutions to the government sector, the revolving door phenomenon is observable and feasible solely between government executive institutions and sovereign legislative bodies. As a result, the study excludes semi-private and private banks from the statistical population. The study thus focuses exclusively on nine state-owned banks: *National Bank, Sepah Bank, Agricultural Bank, Housing Bank, Industry and Mine Bank, Export Development Bank of Iran, Post Bank, Cooperative Development Bank,* and *Mehr Iran Gharz-al-Hasna Bank*.

The dataset consists of 173 bank-year observations, structured as panel data. Following the 2019 merger of five military banks into *Sepah Bank*, the bank's financial statements have yet to be released. Also, *Mehr Iran Gharz-al-Hasna Bank* was established in 2007, and its financial statements were thus excluded from the sample before that year. All data and statistics used in this study were obtained from the financial statements available on the websites of the Central Bank of the Islamic Republic of Iran and the banks.

3.1. The model

After identifying the factors influencing bank performance based on the framework of Athanasoglou et al. (2009) and through its subsequent extension and modification, the primary and revised model employed in this study is presented as follows:

$$ROA_{it} = \beta_0 + \beta_1 CA_{it} + \beta_2 AQ_{it} + \beta_3 MC_{it} + \beta_4 Earnings_{it} + \beta_5 Liq_{it} + \beta_6 CA.RD_{it} + \beta_7 AQ.RD_{it} + \beta_8 MC.RD_{it} + \beta_9 Earnings.RD_{it} + \beta_{10} Liq.RD_{it} + \beta_{11} Inflation_t + \beta_{12} GDP_t + \varepsilon_{it} \quad (1)$$

Where return on assets (ROA) is calculated as net profit divided by total assets for bank i in year t. Capital adequacy ratio (CA) is the ratio of shareholders' equity to total assets of bank i in year t. Asset quality (AQ) shows the ratio of total liabilities to total assets of bank i in year t. Management ratio (MC) defines Operating income divided by total assets for bank i in year t profitability ratio (Earnings) indicates total loans granted/total deposits for bank i in year t; Liquidity Ratio presents the time deposits/totals asset of a bank in year a Time Period Length inflation and gross domestic product (GDP) were used as external determinant affecting number one performances of banks

Additionally, RD_it was included as a variable in the model to investigate the influence of the revolving door effect on bank performance and interaction terms between independent variables (internal determinants) and RD_it. This is a dummy variable taking the value of one if the positions of CEO or board member between the Central Bank of the Islamic Republic of Iran and state-owned banks experience a turnover, zero otherwise (Table 1).

First, return on assets (ROA) is used as a measure for financial performance in the study due to having two influences. First, ROA is recognised in the banking literature as a sound proxy for operational efficiency (see Athanasoglou et al., 2009 for a comprehensive overview), particularly within studies exploring how internal management and governance mechanisms influence this outcome. As a direct measure of the degree to which banks utilize their asset base in pursuing profits, it is also a ratio that assesses the relevance of revolving-door dynamics. Second, reliable and comparable data for Iranian state-owned banks were not consistently available across the study period for alternative measures, such as return on equity (ROE) and market-based indicators (e.g., Tobin's Q and stock returns). This limitation arises from the fact that these banks are not publicly traded and operate under distinctive ownership and reporting structures, which restricts the use of market-based metrics.

Table 1. Summary of model variables

Variable		Abbreviation	Measurement Method	References
Dependent	**Return on Assets Ratio**	***ROA_{it}***	**Net Income/Total Assets**	**Athanasoglou et al. (2009)**
Independent	Capital Adequacy Ratio	CA_{it}	Equity/Total Assets	Lopez (1999)
	Asset Quality Ratio	AQ_{it}	Total Liabilities/Total Assets	Lopez (1999)

	Management Ratio	MC_{it}	Operating Income/Total Assets	Lopez (1999)
	Profitability Ratio	$Earnings_{it}$	Total Loans/Total Deposits	Lopez (1999)
	Liquidity Ratio	Liq_{it}	Deposits/Total Assets	Lopez (1999)
	Revolving-Door Index	RD_{it}	Dummy variable (1 = transfer occurred; 0 = otherwise)	Vanatta (2024) and Chalmers et al. (2022)
Control	Inflation Rate	$Inflation_t$	Annual percentage change in CPI	The Central Bank of the Islamic Republic of Iran (Database)
	Gross Domestic Product Index	GDP_t	Annual growth rate of GDP (%)	The Central Bank of the Islamic Republic of Iran (Database)

4. Findings

The subsequent section presents the research findings, including descriptive statistics, homogeneity and correlation tests, regression analysis, and robustness checks using additional variables to assess the research model.

4.1. Descriptive statistics

Table 2 presents descriptive statistics for the research variables, summarizing each variable's data distribution.

Table 2. Descriptive statistics

Variable	Mean	Median	Max	Min	S. Deviation
ROA	-0.007	0.003	0.101	-0.485	0.072
CA	0.130	0.068	0.854	-0.077	0.152
AQ	0.856	0.934	1.425	0.144	0.200
MC	0.064	0.053	0.599	-0.041	0.058
Earnings	1.383	1.157	3.795	0.007	0.895
Liq	0.307	0.290	0.792	0.003	0.202
CA.RD	0.010	0.000	0.201	-0.051	0.033
AQ.RD	0.138	0.000	1.051	0.000	0.328
MC.RD	0.008	0.000	0.109	0.000	0.021
Earnings.RD	0.209	0.000	3.344	0.000	0.599
Liq.RD	0.045	0.000	0.557	0.000	0.122
Inflation	21.661	15.800	47.100	9.000	12.09
GDP	2.696	3.300	14.163	-8.555	5.130

4.2. Unit Root Test (Stationarity Test)

Ensuring stationarity within the series used in a model is essential to avoid misleading statistical inferences and erroneous regression results. To ensure robust analysis, the stationarity of the variables was examined using the Augmented Dickey–Fuller test for unit roots. A summary of the unit root test results appears in Table 3. To select the optimal lag length, the study applied the Schwarz-Bayesian Criterion.

Table 3. Results of the Unit Root Test

Series	Augmented Dickey–Fuller Test Statistic	Critical Values			Degree	Results
		%1	%5	%10		
ROA	-6.355	-3.469	-2.879	-2.576	I (0)	In levels

CA	-4.521	-3.469	-2.879	-2.576	I (0)	In levels
AQ	-7.370	-3.469	-2.879	-2.576	I (0)	In levels
MC	-6.200	-3.469	-2.879	-2.576	I (0)	In levels
Earnings	-4.831	-3.469	-2.879	-2.576	I (0)	In levels
Liq	-4.344	-3.469	-2.879	-2.576	I (0)	In levels
CA.RD	-8.573	-3.469	-2.879	-2.576	I (0)	In levels
AQ.RD	-4.866	-3.469	-2.879	-2.576	I (0)	In levels
MC.RD	-7.907	-3.469	-2.879	-2.576	I (0)	In levels
Earnings.RD	-8.341	-3.469	-2.879	-2.576	I (0)	In levels
Liq.RD	-7.175	-3.469	-2.879	-2.576	I (0)	In levels
Inflation	-9.764	-3.469	-2.879	-2.576	I (0)	In levels
GDP	-8.023	-3.469	-2.879	-2.576	I (0)	In levels

Based on the results of the Augmented Dickey–Fuller unit root test in Table 3, all the research variables are in levels and do not exhibit a unit root. Consequently, the regression outcomes are not spurious.

4.3. Correlation coefficient results

Table 4 displays the correlation coefficients between the dependent variable and the independent variables. These results illustrate the relationship between the return on assets and the independent variables in the study and indicate the degree of dispersion around the regression line.

Table 4. Correlation coefficient results

	ROA	*CA*	*AQ*	*MC*	*Earnings*	*Liq*	*CA.RD*	*AQ.RD*	*MC.RD*	*Earnings.RD*	*Liq.RD*	*Inflation*	*GDP*
ROA	1	0.083	-0.076	0.357	-0.350	0.043	-0.023	-0.107	-0.072	-0.099	-0.102	-0.039	0.019
CA		1	-0.627	0.132	0.292	-0.236	0.016	-0.137	-0.088	-0.088	-0.122	-0.133	0.193
AQ			1	0.010	-0.392	0.256	-0.079	0.048	0.006	-0.047	0.015	0.096	-0.034
MC				1	-0.339	0.173	-0.024	-0.060	0.012	-0.071	-0.071	0.014	-0.048
Earnings					1	-0.567	0.008	-0.079	-0.066	0.040	-0.081	0.081	-0.044
Liq						1	-0.063	0.012	-0.024	-0.080	0.036	-0.092	-0.023
CA.RD							1	0.744	0.849	0.863	0.801	0.098	-0.015
AQ.RD								1	0.949	0.912	0.967	-0.039	0.024
MC.RD									1	0.925	0.936	0.015	0.011
Earnings.RD										1	0.908	0.072	-0.027
Liq.RD											1	-0.039	0.009
Inflation												1	-0.508
GDP													1

Based on the correlation coefficients in Table 4, a negative relationship exists between the dependent variable (return on assets) and the independent variables, including the asset quality ratio, profitability ratio, and inflation index. Consequently, the findings suggest that increases in these variables are associated with decreases in the return on assets. Furthermore, the correlation between these variables and the return on assets is weak, indicating relatively high dispersion around the regression line.

In contrast, the results indicate a positive relationship between the dependent variable and the following independent variables, including the capital adequacy ratio, management efficiency ratio, liquidity ratio, and the gross domestic product index (with correlation coefficients of 0.083, 0.357, 0.043, and 0.018, respectively). These coefficients indicate that increases in these variables are associated with higher asset returns for state-owned banks. Moreover, the correlation between the management efficiency ratio and the dependent variable exhibits minimal dispersion around the regression line.

4.4. Regression estimation

To test the research model, the results of the multivariate regression, along with the t-test, F-test, and Durbin-Watson test, are presented in Table 5. The results were analyzed at the 5% significance level.

Table 5. Regression estimation results

$$ROA_{it} = \beta_0 + \beta_1 CA_{it} + \beta_2 AQ_{it} + \beta_3 MC_{it} + \beta_4 Earnings_{it} + \beta_5 Liq_{it} + \beta_6 CA.RD_{it} + \beta_7 AQ.RD_{it} + \beta_8 MC.RD_{it} + \beta_9 Earnings.RD_{it} + \beta_{10} Liq.RD_{it} + \beta_{11} Inflation_t + \beta_{12} GDP_t + \varepsilon_{it}$$

Component	Coef.	S. deviation	t-statistic	p-value
intercept	2.934	0.800	3.665	0.000
CA	-0.182	0.886	-0.206	0.837
AQ	-1.431	0.661	-2.167	0.032
MC	5.455	1.861	2.931	0.004
Earnings	-0.730	0.144	-5.061	0.000
Liq	-1.429	0.573	-2.496	0.014
CA.RD	1.447	1.220	1.186	0.238
AQ.RD	0.915	2.747	0.333	0.739
MC.RD	-25.219	31.248	-0.807	0.421
Earnings.RD	0.019	0.891	0.022	0.983
Liq.RD	-3.595	5.715	-0.629	0.530
Inflation	-0.004	0.009	-0.386	0.670
GDP	-0.003	0.022	-0.148	0.883
F-statistic		4.835		
p-value		0.000		
Adjusted R-squared		0.280		
Adjusted R-squared		0.222		
Durbin–Watson Statistic		1.011		

Based on the results in Table 5, the coefficient for the asset quality ratio (AQ) is -1.431, which indicates that a one-percentage-point increase in the asset quality ratio is associated with a 1.431-percentage-point decrease in the return on assets (ROA). At the 5% significance level, the p-value for this coefficient is 0.032, suggesting a statistically significant relationship between asset quality and ROA.

Similarly, the coefficient for the management efficiency ratio (MC) is 5.455, implying that a 1-percentage-point increase in management efficiency corresponds to a 5.455-percentage-point increase in ROA. The p-value for this coefficient is 0.004, confirming a significant positive relationship between management efficiency and ROA at the 5% significance level.

Furthermore, the coefficients for the profitability ratio (Earnings) and the liquidity ratio (Liq) are -0.730 and -1.429, respectively. These negative coefficients suggest that increases in profitability and liquidity ratios are associated with decreases in ROA. The p-values for these coefficients are 0.000 and 0.014, respectively, indicating statistically significant negative relationships at the 5% significance level.

Conversely, other independent variables, including the capital adequacy ratio (CA), inflation rate (Inflation), gross domestic product (GDP), and the interaction terms involving the revolving-door dummy variable (CA.RD, AQ.RD, MC.RD, Earnings. RD, and Liq.RD), do not exhibit statistically significant relationships with ROA.

The overall multivariate regression model is statistically significant, as evidenced by an F-test p-value of 0.000, which is below the 0.05 threshold. With an R-squared value of 0.28, the model explains roughly 28% of the variation in ROA through its independent variables.

Additionally, the Durbin-Watson statistic reported in Table 5 is slightly above one, indicating no autocorrelation in the residuals. This result shows that the residuals follow a random pattern, confirming that the regression assumption of independence holds.

4.5. Redundant variables test

The Redundant Variables Test tests the hypothesis that a set of variables in a regression equation may be unnecessary. Specifically, this test assesses whether the coefficients of these variables are jointly equal to zero, implying that the model may omit them without compromising its explanatory power. In this context, the null hypothesis posits that the interaction terms between the dummy variable representing managerial transitions (RD) and the independent variables: capital adequacy (CA.RD), asset quality (AQ.RD), management ratio (MC.RD), earnings ratio (Earnings.RD), and liquidity ratio (Liq.RD) are redundant in the multiple regression model. Table 6 presents the results of the Redundant Variables Test.

Table 6. Results of the redundant variables test

Null Hypothesis	F-Statistic	Log-Likelihood Ratio	p-value (F-Statistic)	p-value (Chi-Square)	Test Result
The interaction terms (*CA.RD, AQ.RD, MC.RD, Earnings.RD, Liq.RD*) are redundant in the regression model.	0.977	5.225	0.434	0.389	Null Hypothesis Accepted

As indicated in Table 6, the p-value associated with the F-statistic is 0.434, which exceeds the conventional significance level of 0.05. Consequently, we fail to reject the null hypothesis, suggesting that the interaction terms (CA.RD, AQ.RD, MC.RD, Earnings. RD, Liq.RD) do not contribute significantly to the model and can be considered redundant.

5. Conclusions and discussions

This study employed the model developed by Athanasoglou et al. (2009) to examine the impact of the revolving door phenomenon on the performance of state-owned banks in Iran. By incorporating a dummy variable representing the transfer and rotation of executives and board members between the Central Bank of the Islamic Republic of Iran and state-owned banks, the study aimed to assess the influence of such movements on bank performance.

The findings indicated that none of the independent variables, adjusted for managerial transitions, exhibit a statistically significant relationship with Return on Assets (ROA). This lack of significance suggests that the revolving door phenomenon and the potential conflicts of interest it may entail do not materially affect the profitability of state-owned banks in Iran.

As previously discussed, conflicts of interest arising from the movement of executives between the Central Bank and state-owned banks can lead to corruption or rent-seeking behaviors. However, the results of this study imply that managerial transitions linked to the revolving door phenomenon have not adversely affected the financial performance (ROA) of state-owned banks. The results suggest that executives transitioning between these institutions maintain integrity and do not engage in conflicts of interest or use their positions for personal advantage. Furthermore, stringent oversight by relevant authorities during their tenure mitigates potential conflicts of interest, corruption, and rent-seeking activities.

Hence, the lack of a statistically significant relationship between revolving door practices and the financial performance of Iranian state-owned banks can be explained by several factors. One possible explanation lies in the institutional environment: Formalized procedures and supervisory mechanisms regulate managerial transfers between the Central Bank and state-owned banks, thereby limiting opportunities for conflicts of interest or regulatory capture. In such a setting, revolving door movements may have a neutral rather than a decisive impact on performance outcomes.

At the same time, the study should explicitly acknowledge methodological considerations. The study's methodology measures revolving-door practices using a dummy variable that captures managerial transfers but excludes their duration, frequency, or intensity. Moreover, the exclusive reliance on ROA as the performance indicator, while consistent with the banking literature, may not fully capture market-based, reputational, or long-term governance effects. These factors may have contributed to the absence of statistical significance in the results.

Based on these findings, future research should explore conflicts of interest and the revolving door phenomenon in post-employment contexts, particularly before and after retirement. The lack of post-retirement oversight may facilitate the transfer of occupational rents to subsequent employment phases, potentially leading to conflicts of interest.

The findings differed in several key aspects from those reported in studies conducted in developed economies, where institutional and regulatory structures are stronger. Empirical studies in advanced economies such as the United States and Europe have often linked revolving-door practices to positive outcomes, including improved stock market performance and better risk management (Shive and Forster, 2017; Canayaz et al., 2015). Previous studies have typically attributed these results to the transfer of regulatory expertise and privileged access to policy information. By contrast, our analysis of Iranian state-owned banks finds no statistically significant effect. This suggests that under conditions of strong state control and formal oversight, the revolving door may not significantly influence financial performance.

The results also stood in contrast to evidence from China (Li, 2023), where revolving door appointments were associated with declines in profitability and financing constraints due to perceptions of political entanglement. In the Iranian case, the absence of a significant relationship pointed instead to a neutral outcome that revolving door movements neither generate measurable efficiency gains nor impose discernible financial costs. This comparison highlights the critical role of institutional context in shaping the outcomes of revolving door practices.

Taken together, these contrasts reinforce the view that the revolving door phenomenon is not universal in its effects but is instead contingent on governance structures, ownership patterns, and regulatory environments. This study contributes to broadening the global debate by documenting a neutral impact in the Iranian state-owned banking sector, which highlights the need for more context-sensitive approaches in analyzing the intersection of governance, regulation, and bank performance.

In addition, the limited number of observations in the multivariate regression analysis posed a constraint in this study. Expanding the sample size in future research could address this limitation and enhance the robustness of the findings.

Future research should expand the analysis of revolving-door dynamics in state-owned banks by addressing multiple dimensions. First, studies could investigate managerial transitions at levels below the CEO and board, as well as post-employment movements, to capture the full scope and intensity of personnel rotations. Second, qualitative approaches, including case studies, interviews, and archival analyses, are needed to assess non-financial consequences such as distortions in credit allocation, regulatory leniency, factional influence, and impacts on public trust. Finally, research should incorporate a broader set of performance and governance indicators beyond accounting profitability, such as non-performing loans, loan quality, regulatory enforcement, and transparency metrics, enabling a more comprehensive understanding of how revolving-door practices affect institutional integrity and the stability of state banking systems.

References

1. Athanasoglou, P. P., Georgiou, E. A. and Staikouras, C. C. (2009). Assessing output and productivity growth in the banking industry. *The Quarterly Review of Economics and Finance*, 49(4), pp. 1317-1340. https://doi.org/10.1016/j.qref.2009.08.003
2. Barth, A. and Mansouri, S. (2021). Corporate culture and banking. *Journal of Economic Behavior & Organization,* 186(1), pp. 46-75. https://doi.org/10.1016/j.jebo.2021.02.010
3. Blanes i Vidal, J., Draca, M. and Fons-Rosen, C. (2012). Revolving door lobbyists. *American Economic Review,* 102(7), pp. 3731-3748. https://doi.org/10.1257/aer.102.7.3731
4. Canayaz, M., Martinez, J. V. and Ozsoylev, H. N. (2019). Is the revolving door of Washington a back door to government contracts and excess returns?. Pennsylvania State University, Pennsylvania. Available at SSRN 2596520. https://doi.org/10.2139/ssrn.2596520
5. Chalmers, A. W., Klingler-Vidra, R., Puglisi, A. and Remke, L. (2022). In and out of revolving doors in European Union financial regulatory authorities. *Regulation & Governance*, 16(4), pp. 1233-1249. https://doi.org/10.1111/rego.12424
6. Esmaeilzadeh Mogheri, A. (2009). The survey on inflation effectiveness from total investment in Iran's economy. *Journal of Economic Research*, 9(2), pp. 97–123 (In Persian). https://joer.atu.ac.ir/article_2899_749b7fcf843b2674985b738ebac652e8.pdf?lang=en
7. Goddard, J., Molyneux, P. and Wilson, J. O. (2004). Dynamics of growth and profitability in banking. *Journal of Money, Credit and Banking*, 36(6), pp. 1069-1090. https://doi.org/10.1353/mcb.2005.0015.
8. Li, Z. (2023). Connections as liabilities: The cost of the politics–business revolving door in China. *British Journal of Political Science,* 53(4), pp. 1252-1272. https://doi.org/10.1017/S0007123422000473.
9. Lopez, Jose A. (1999). Using CAMELS ratings to monitor bank conditions. *FRBSF Economic Letter 1999-19* (June 11). https://www.frbsf.org/research-and-insights/publications/economic-letter/1999/06/using-camels-ratings-to-monitor-bank-conditions/
10. Lucca, D., Seru, A. and Trebbi, F. (2014). The revolving door and worker flows in banking regulation. Journal of Monetary Economics, 65, pp. 17-32. https://doi.org/10.1016/j.jmoneco.2014.05.005.
11. Mehrbanpour, M., Naderi Noorain, M., Inanlou, E. and Ashari, A. (2017). Factors influencing bank profitability. *Empirical Studies in Financial Accounting*, 14(54), pp. 119–140 (In Persian). https://doi.org/10.22054/qjma.2018.8272.
12. Menon, K. and Williams, D. D. (2004). Former audit partners and abnormal accruals. *The Accounting Review*, 79(4), pp. 1095-1118. https://doi.org/10.2308/accr.2004.79.4.1095
13. Monsef, A. and Mansouri, N. (2011). The study of effective factors on the level of bank deposits (emphasizing on the rate of profit-sharing securities 1988 - 2008). *Journal of Monetary and Financial Economics,* 17(34), pp. 69-91. (In Persian). https://doi.org/10.22067/pm.v17i34.27351.
14. OECD (2004), Managing Conflict of Interest in the Public Service: OECD Guidelines and Country Experiences, OECD Publishing, Paris, https://doi.org/10.1787/9789264104938-en
15. Roudini, S. (2023). Unlocking the revolving door: How FDA-firm relationships affect drug approval rates and innovation in the pharmaceutical industry. Social Science Research Network, University of Memphis, Memphis, Tennessee. Available at SSRN, http://dx.doi.org/10.2139/ssrn.4658800
16. Shive, S. A. and Forster, M. M. (2017). The revolving door for financial regulators. *Review of Finance*, 21(4), pp. 1445-1484. https://doi.org/10.1093/rof/rfv044

17. Talebi, M. and Salgi, M. (2016). Examining the relationship between risk and capital adequacy ratio: evidence from Iranian banks. *Monetary and Banking Research Quarterly*, 9(30), pp. 513–543 (In Persian). https://dor.isc.ac/dor/20.1001.1.26453355.1395.9.30.5.0.
18. Tenekedjieva, A. M. (2020). The revolving door and insurance solvency regulation. The University of Chicago, Chicago, Illinois. https://doi.org/10.2139/ssrn.3762573
19. Vanatta, Sean H. (2024). Revolving door governance: bank supervisors in the United States, 1863–1933. *Management and Organizational History*, 19(3), pp. 153-181. https://doi.org/10.1080/17449359.2024.2353797.
20. Williams, John C. (2017). The global growth slump: causes and consequences. *FRBSF Economic Letter 2017-19* (July 3). https://www.frbsf.org/research-and-insights/publications/economic-letter/2017/07/global-growth-slump-causes-consequences-speech/

IJAAF; Vol. 10, No. 2, Spring 2026, Serial 35

Ferdowsi University of Mashhad

Iranian Journal of
Accounting, Auditing & Finance

Quarterly

RESEARCH ARTICLE

Predicting Stock Market Returns Using Temporal Fusion Transformer: A Comprehensive Data-Driven Approach

Somayeh Mohebi*
School of Industrial Engineering and Management, Shahrood University of Technology, Shahrood, Iran
Mohammad Reza Mohebbi
Department of Computer Science, University of Passau, Passau, Germany
Javad Mohebbi Najm Abad
Department of Computer Engineering, Imam Reza International University, Mashhad, Iran

How to cite this article:
Mohebi, S., Mohebi, M. R. and Mohebi Najm Abad, J. (2026). Predicting Stock Market Returns Using Temporal Fusion Transformer: A Comprehensive Data-Driven Approach. Iranian Journal of Accounting, Auditing and Finance, 10(2), 125-151. doi: 10.22067/ijaaf.2026.47513.1562
https://ijaaf.um.ac.ir/article_47513.html

ARTICLE INFO

Article History
Received: 2025-06-14
Accepted: 2025-12-06
Published online: 2026-04-12

Keywords:
Deep Learning, Dimensionality Reduction, Feature Selection, Stock Market Prediction, Temporal Fusion Transformer

Abstract

Stock market return prediction remains a highly challenging task due to the complex, dynamic, and noisy nature of financial markets. Although machine learning and deep learning models have been widely explored, many approaches struggle to capture long-term temporal dependencies and provide interpretable feature selection. In tackling these issues, the Temporal Fusion Transformer (TFT) is applied for multi-horizon time series forecasting of daily returns in the Tehran Stock Exchange (TSE). Accordingly, the study proposes a new feature selection method called Initial Selected Features–Mutual Information Difference (ISF-MID), which enhances the traditional Minimum Redundancy Maximum Relevance (mRMR) algorithm by demonstrating a superior capability of handling redundancy and focusing on determining important parameters. Dual preprocessing works on ISF-MID, mRMR and PCA that improve the predictive performance, while also allowing for interpretable machine learning by working with a better representation of input. Mean Absolute Error (MAE), Mean Squared Error (MSE) and the coefficient of determination (R^2) were used to conduct extensive comparisons with benchmark methods such as Long Short-Term Memory LSTM, Multi-Layer Perceptron MLP, and Random Forest RF. Thus, the proposed method achieved competitive performance against benchmark architectures (i.e. TFT achieves R^2 of 98.9%, MAE 0.00043 and MSE 0.000018 on out-of-sample test, as well as high directional accuracy). Overall, the synergy of TFT with the ISF-MID feature selection strategy provides methodological novelty and practical significance, extending a potential framework for evidence-based return prediction and informed risk-central investment decisions.

https://doi.org/10.22067/ijaaf.2026.47513.1562

CC BY

NUMBER OF REFERENCES
45

NUMBER OF FIGURES
10

NUMBER OF TABLES
7

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

*Corresponding Author: Somayeh Mohebi
Email: mhd_moradi@um.ac.ir
Tel: 09153215094
ORCID: 0000-0003-2559-7440

1. Introduction

The prediction of market returns is probably the essential and extremely challenging issue in financial analysis, as markets are mostly dynamic, nonlinear and noisy—especially in real-world data for time series (Sharma & Mehta, 2024). Generating accurate forecasts helps with portfolio management, risk control, and investment strategy design. However, it is still challenging because of the complexity and high dimensionality of these influencing factors. Interactions among macroeconomic factors, industry dynamics, company performance, and investor behavior are difficult to capture with reliable models (Dos Santos et al., 2025). These difficulties have led to considerable research on predictive techniques of some kind to identify regularities in complex phenomena of financial data (Sharma and Mehta, 2024).

Conventional Machine Learning (ML) and Deep Learning (DL) methods have been extensively studied as part of financial forecasting. Support Vector Regression (SVR), Random Forest (RF), and recurrent neural networks such as Long Short-Term Memory (LSTM) have demonstrated excellent capability to model certain nonlinearities of financial time series despite essential drawbacks (Tulcanaza-Prieto and Lee, 2019). Recurrent structures are typically unable to describe long-term relationships, ensemble and kernel approaches are hard to interpret, and most of them are still vulnerable to irrelevant or redundant variables, which may impair prediction precision at high dimensionalities of datasets (Alam et al., 2024).

The latest progress of sequence modeling, such as transformer architecture, provides the potential to overcome these limitations. The Temporal Fusion Transformer (TFT) is arguably one of the best DL models introduced so far, designed for multi-horizon forecasting (Lim et al., 2021). The use of attention mechanisms allows the model itself to automatically assign flexible importance to input features over time, making it particularly efficient under conditions of temporal complexity and signal noise (Hartanto and Gunawan, 2024). Although other applications have flourished with transformers, the usage of the transformers applied to financial return prediction is limited. Specifically, limited studies have concentrated on the integration of TFT with systematic feature engineering that has been tailored based on the unique features of financial markets (Davoud Abadi, 2025).

Feature selection and dimension reduction are known to be important factors contributing to the performance of models in financial predictions. The models are more accurate and by eliminating redundancy, they are much more interpretable. Minimum Redundancy Maximum Relevance (mRMR) and Principal Component Analysis (PCA) are typical methods that are widely adopted to improve data quality before model training (Mohebi et al., 2022a). Although most of these methods are useful, they are usually used in isolation, and only a few studies have investigated hybrid pipelines, combining feature selection and feature extraction. Additionally, existing work has rarely evaluated the establishment of new feature selection methods that are explicitly designed for financial time series data.

This study proposes a forecasting framework, filling these gaps by introducing a powerful framework that predicts the daily returns of the Tehran Stock Exchange (TSE). Accordingly, the Initial Selected Features–Mutual Information Difference (ISF-MID) approach is proposed as a newly developed feature selection framework, which generalizes the mRMR method to enhance relevance and reduce redundancy. ISF-MID is used as a preprocessing technique to improve prediction accuracy and interpretability, with its performance also compared to that of PCA. The TFT is used as the main forecasting framework and compared with several ML and DL baselines, such as Multi-Layer Perceptron (MLP), Radial Basis Function (RBF), SVR, Decision Trees (DT), RF, Deep Neural Networks (DNN), and LSTM. The overall research pipeline is illustrated in Figure 1, covering dataset construction, preprocessing, feature selection and extraction, and forecasting with TFT and comparative baselines. The next sections describe each step in detail. The principal contributions of this paper are as follows:

We developed a rich set of candidate features at scale by combining market indices, technical indicators and macroeconomic variables. Based on this, secondary features were produced that enhance the primary input by adding more dependencies to create a richer representation of the market behavior. We provided a detailed comparison of the performance of the TFT with popular classical ML and modern DL models for daily stock return forecasting in the TSE. We proposed a new feature selection approach termed "ISF-MID", which utilizes the mRMR principle of maximizing relevance while reducing redundancy, leading to compact and efficient input representation. We compared alternative preprocessing strategies, including an mRMR method and its two variants, MID and FCD, as well as a PCA strategy to assess their relative impact on predictive performance, establishing that the ISF-MID variant is a more effective extension of mRMR. We showed a TSE index prediction case study that targeted feature selection framework ranks and validates the most impactful features for predicting the TSE index. In using the proposed ISF-MID algorithm, this pipeline pushes methodological innovation whilst giving more interpretable financial forecasting.

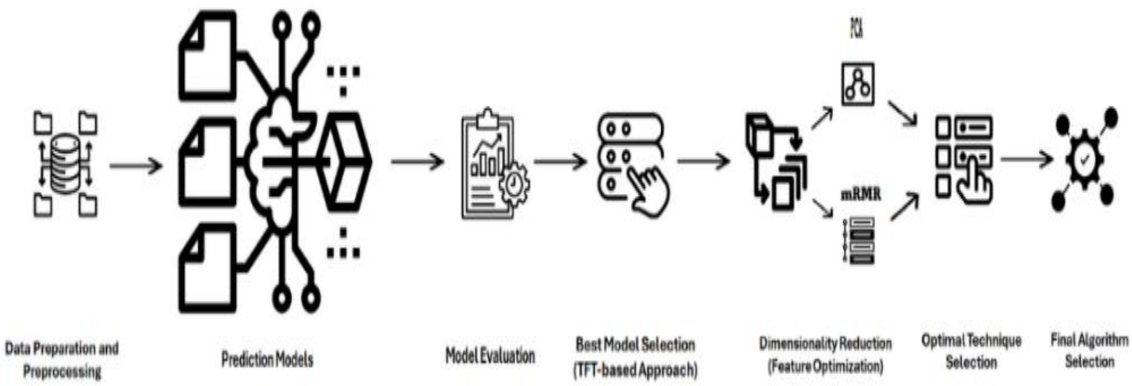

Figure 1. Proposed framework for stock index return prediction

2. Related work

This prediction of stock market returns is a very difficult problem because financial data is inherently dynamic, nonlinear and noisy (Zheng et al., 2024). Proper prediction assists investment decisions, risk management, and portfolio allocation. Reported literature contains numerous approaches from typical ML models to more complex DL, hybrid forms, RL and transformer-based methods (Li et al., 2020; Sezer et al., 2020). However, there still exist limitations in these studies regarding the ability to learn long-term dependencies, interpret feature importance and high dimensionality. To demonstrate the overview in a structured manner, we grouped related work into three categories: (i) ML and DL models; (ii) hybrid and RL approaches; and (iii) transformer-based frameworks focused on more recent TFT applications.

2.1. Machine learning and deep learning models

Traditional ML methods have been popular in stock market prediction and are also able to capture structured and high-dimensional financial data. As a result, SVR, DT, RF, KNN (k-Nearest

Neighbors) and GBM (Gradient Boosting Machines) are all well proven to have the ability to model short-term relationships and cross-sectional appearances in financial indicators (Kumbure et al., 2022). Feature engineering and dimensionality reduction are often key complements: e.g., Wei et al. (2021) found that PCA reduces the redundancies of these features, retaining important information in the market. We have also proven a useful ensemble-based model; the adaptation of AdaBoost with DTs is associated with increasing predictive precision and computational efficiency (Guo et al., 2017), while Zhang et al. (2018) reduced complexity without sacrificing accuracy using combined feature engineering with ensemble learning strategies. While some domains were successful, such methods tend to fall short of adequately learning the complex sequential and nonlinear structures present within financial time series.

Simultaneously, DL models have become a popular tool to handle temporal dependencies and nonlinear relationships in financial data. Their effectiveness, however, is limited mainly due to the presence of long-term dependencies, upon which LSTM networks and Gated Recurrent Units (GRU) were developed as recurrent architectures that are widely used. Nabipour et al. (2020) showed that LSTM outperformed classical ML algorithm classes for trend forecasting at the TSE consistently, while Wei et al. (2024) found the same thing. Unlike the recurrent models, other data representations, such as the use of bidirectional networks and autoencoder based architectures, have been proposed to capture complex feature interactions and latent structures (Gandhudi et al., 2024; Bao et al., 2017). Moreover, where inputs were often restricted to numbers, adding textual sentiment from news and social networks transformed the forecast in DL frames. This was accomplished by combining LSTM or MLP networks with natural language processing modules (Awad et al., 2023).

Yet the ML and DL approaches suffer from significant caveats. Traditional ML models usually do not consider temporal dynamics. Although DL architectures are more powerful but are subject to overfitting, need high computation resources, and they also behave like “black boxes” as they have less interpretability. Additionally, the two model classes remain fundamentally sensitive to redundant or irrelevant variables, a very relevant issue in high-dimensional financial data sets. These limitations underscore the need for systematic feature selection and/or dimensionality reduction to improve predictive robustness as well as interpretability. As a result, more recent work has focused on sophisticated architectures, and in particular transformer-based models, that are not only able to capture long-term dependencies but can also be enhanced with principled feature engineering strategies.

2.2. Hybrid and reinforcement learning approaches

A number of research-based frameworks have been proposed that integrate the complementary paradigms of stand-alone ML and DL models to mitigate their individual drawbacks, thus improving forecasting performance. These methods generally combine statistical preprocessing and deep sequence models or optimization techniques with non-linear predictors. For example, Wu et al. included the work of Yi and Shen (2023), who queried historical stock data using Triple Q-learning in conjunction with the Growing Neural Gas (GNG) algorithm. John and Latha (2023) proposed a Hybrid Recurrent Neural Network (HyRNN), which incorporated financial news sentiment into BiLSTM and GRU. Feature decomposition methods have also been widely applied. Liang et al. (2019) adopted Wavelet Transform (WT) and LSTM for the separation of informative frequency components, whereas Niu et al. (2020) utilized VMD, presenting more stable prediction results. Hybridization has also found its way into optimization-based approaches, with an instant example being the Bayesian optimization carcass algorithm (Liwei et al., 2021). SVR mollusk with Particle Swarm Optimization (PSO)-based hyperlloid-adding model (Jaiwang and Jeatrakul, 2018) allows for on-the-fly hyperparameter changes appropriate for ever-fluctuating market environments.

RL is this brand new, promising direction for the entire world of financial forecasting and algorithmic trading that goes beyond Hybrids. Deep Q-Networks (DQN) and variants with attention features and sentiment features have shown a lot of achievement in navigating SEO Bid environments that are dynamic and non-deterministic (Awad et al., 2023). Recent works leverage multimodal signals in RL to train models on numerical and linguistic data streams together for improved strategy learning. Concurrently, federated learning has proposed frameworks to improve privacy and mitigate data diversity issues. Pourroostaei Ardakani et al. (2023) showed RF and SVM models improved by federated learning to protect privacy across institutions, while Kumarappan et al. (2024) introduced LSTM models integrated with federated learning, balancing superior predictive effectiveness with stringent privacy safeguards. These approaches highlight the adaptability of RL and federated settings to practical financial environments where both robustness and security are critical.

Although hybrid RL and federated approaches represent important advances, they also introduce added architectural complexity and often shift the focus to trading strategy optimization rather than systematic feature engineering. In their dependence upon black box deep structures, they sacrifice interpretability and often end up with redundant or irrelevant features in the input space, undermining robustness. Such gaps drive the investigation of models, which can maintain predictive power, interpretable features and high feature efficiency. The current work bridges this gap by integrating the interpretability and temporal modeling capabilities of TFT with state-of-the-art feature selection and dimensionality reduction techniques to deliver accurate yet interpretable financial forecasts.

2.3. Transformer-based models for financial forecasting

Recently, transformer architectures have been applied in financial forecasting due to their capabilities of capturing long-range dependencies and adaptively weighting relevant signals. To make them work for high-frequency trading data, sparse or decomposition-based attention has been used (Wu, 2023) in informer and outformer. They proposed transformer layers were adapted to multivariate financial series, resulting in improvements over recurrent models (Wang et al., 2022). Emami Gohari et al. (2024) proposed a multimodal transformer architecture that jointly processes numeric time-series and textual inputs (e.g., economic reports) through intra-, inter- and target-modal attention layers, demonstrating improved performance over baseline models in a financial forecasting setting. In addition to numeric features, other models like FinSentiBERT and multimodal transformers demonstrated the ability of textual sentiment (in particular from headlines/analyst reports) to extract predictive signals (Sun et al., 2025). The results of these studies support the acclaimed usefulness of the transformer structure in the main financial features.

More recently, researchers have applied the TFT specifically to financial markets. Hartanto and Gunawan (2024) applied TFT to Indonesian stock prices and found gains over baseline transformer variants, as well as naïve methods in volatility-sensitive settings (consult Hartanto and Gunawan (2024) for datasets, baselines, and metrics). Davoud Abadi (2025) conducted a comparison between TFT and LSTM, SMA combined with SVR on both the S&P 500 and Tehran indices, and reported that for the Tehran index, TFT outperformed the results. Yang et al. (2025) moreover extended TFT into a multi-sensor framework for Sharpe Ratio (SR=expected return/risk) optimization in what is called here as the Tuning Framework Transfer for ASset Risk Optimization (or TFT-ASRO), gaining an 18% risk-adjusted forecasting performance. Dashtaki et al. (2025) developed a cross-modal temporal fusion framework that integrates textual and numerical features, leading to enhanced inference accuracy and efficiency. Lim et al. (2021) discovered that TFT is interpretable in forecasting tasks and validated the transparency of TFT compared to other black-box models.

Although TFT has shown great promise for financial forecasting, existing works typically use raw or less engineered inputs like prices, volumes and sentiment embeddings, which opens the models to redundant features, noise, and has lower robustness in high-dimensional settings. To the best of our knowledge, no prior work systematically integrates TFT with a feature optimization framework in which selection methods (mRMR, ISF-MID) and extraction techniques (PCA) are systematically evaluated as alternative preprocessing strategies. This study provides a novel approach to overcome this by proposing ISF-MID for redundancy-sensitive feature selection and investigating its improvement on both TFT performance measures, predictive accuracy and interpretability with respect to stock index prediction.

3. Methodological background

This part briefly outlines the theoretical underpinnings deployed in this study. We first introduce the Temporal Fusion Transformer (TFT), a state-of-the-art interpretable multi-horizon forecasting model that can utilize both temporal and static covariates. By focusing on the proposed ISF-MID and well-established mRMR and PCA techniques, we then discuss feature selection and dimensionality reduction methods suited for high-dimensional, correlated financial inputs. In combination, these components form the foundation of our forecasting framework.

3.1. Temporal fusion transformer (TFT)

The TFT is a deep learning model, which focuses on interpretable multi-horizon time series forecasting (Lim et al., 2021) and generalizes well with heterogenous datasets containing temporal and static covariates, thus allowing for financial settings in which the underlying dynamics of any market vary according to colliding technical and macroeconomic factors (Lim et al., 2021).

3.1.1. Input representation and embedding

Temporal covariates $\mathrm{x}_t \in \mathbb{R}^{d_t}$ at time t and static covariates $\mathrm{x}^{\text{static}} \in \mathbb{R}^{d_s}$ are mapped to dense embeddings via learnable functions $f_{\text{embed}}^{(t)}(\cdot)$ and $f_{\text{embed}}^{(s)}(\cdot)$:

$$\mathrm{z}_t = f_{\text{embed}}^{(t)}(\mathrm{x}_t), \quad \mathrm{z}^{\text{static}} = f_{\text{embed}}^{(s)}\left(\mathrm{x}^{\text{static}}\right), \tag{1}$$

where $\mathrm{z}_t \in \mathbb{R}^{p_t}$ and $\mathrm{z}^{\text{static}} \in \mathbb{R}^{p_s}$ are latent representations. In financial datasets, this places indicators (e.g., moving averages, exchange rates) in a consistent latent space.

3.1.2. Variable selection networks (VSNs)

VSNs select informative variables at each step [3]. Let $\mathrm{x}_t \in \mathbb{R}^{d_t}$ be the (possibly concatenated) temporal features at time t. The network produces a *weight vector* $\mathrm{w}_t \in \mathbb{R}^{d_t}$: and the bias vector b:

$$\mathrm{w}_t = \mathrm{softmax}(\mathrm{W}\mathrm{x}_t + \mathrm{b}), \tag{2}$$

where $\mathrm{W} \in \mathbb{R}^{d_t \times d_t}$ and $\mathrm{b} \in \mathbb{R}^{d_t}$ are learnable parameters. The i-th entry $[\mathrm{w}_t]_i$ indicates the importance of variable i at time t, improving interpretability by ranking indicators (e.g., volume, oil price).

3.1.3. Gating and residual connections

Gating mechanisms regulate information flow and stabilize training (Hartanto and Gunawan, 2024):

$$\mathrm{h}_t = f_{\text{gate}}\left(\mathrm{z}_t, \mathrm{z}^{\text{static}}\right), \tag{3}$$

where z_t is the main input representation at time step t. Also, z^{static} is time-invariant information. Finally, $f_{\text{gate}}(0)$ denotes a learned gating function. Gates suppress irrelevant/noisy inputs, while residual connections maintain gradient flow.

3.1.4. Multi-head self-attention

Long-range temporal dependencies are modeled via multi-head self-attention (Lim et al., 2021):

$$\text{Attention}(Q, K, V) = \text{softmax}\left(\frac{QK^{\top}}{\sqrt{d_k}}\right)V, \tag{4}$$

with query Q, key K, value V, and key dimension d_k. This uncovers lagged effects (e.g., macroeconomic shocks).

3.1.5. Decoder and forecast generation

Context vectors are decoded to produce multi-horizon forecasts. TFT yields both accurate predictions and post-hoc interpretability via attention and variable selection, attributing importance to covariates and temporal patterns (Lim et al., 2021).

Transition.

Given the high dimensionality and correlation structure of financial inputs, we next outline feature selection and dimensionality reduction methods used to construct compact, informative representations.

3.2. Feature selection and dimensionality reduction

Financial datasets often include redundant variables that can impair generalization. We therefore consider ISF-MID and mRMR (feature selection) and PCA (feature extraction) to improve efficiency and interpretability.

3.2.1. Initial selected features–mutual information difference

Initial Selected Features–Mutual Information Difference (ISF-MID) extends the mRMR-MID criterion by seeding the selection with a domain-informed initial set S_0 and then iteratively adding features from the candidate pool F (the full set of candidate features) based on a redundancy-aware score (Mohebi et al., 2022b). Let S denote the current selected subset (initialized as S_0), y the target variable, and $I(0; 0)$ mutual information. At each step, f^*, which is the next feature to select, is calculated by the following equation:

$$f^* = \arg\max_{f \in F \setminus S} \left[I(f; y) - \alpha \cdot \frac{1}{|S|} \sum_{s \in S} I(f; s) \right], \tag{5}$$

where $\alpha \geq 0$ controls redundancy penalization. $F \setminus S$ are the remaining features, which are not yet selected. Seeding with S_0 injects expert knowledge, the MI term favors relevance, and the penalty discourages selecting near-duplicate indicators (e.g., overlapping moving averages).

In contrast to ISF-MID, which emphasizes redundancy-sensitive feature selection, the traditional mRMR method has been widely applied as a baseline for balancing feature relevance and redundancy.

3.2.2. Minimum redundancy maximum relevance (mRMR)

According to Mohebbi et al. (2025(a)), the mRMR goal strikes a balance between y-relevant information and pairwise redundancy within S:

$$\max_{S\subseteq F,\,|S|=k}\left[\frac{1}{|S|}\sum_{f\in S} I\,(f;y)\;-\;\frac{1}{|S|^2}\sum_{f_1,f_2\in S} I\,(f_1;f_2)\right]. \qquad (6)$$

Here, F is the full feature set and k is the desired subset size. The result is a predictive yet diverse set of variables.

Dimensionality reduction techniques, such as principal component analysis (PCA), offer an alternative to feature selection methods like ISF-MID and mRMR by converting features into orthogonal components, which preserve variance while minimizing redundancy.

3.2.3. Principal component analysis (PCA)

Orthogonal components that account for the most variance are projected by PCA from correlated inputs (Pourroostaei Ardakani et al., 2023). Let Σ be the covariance of centered data, with eigenpairs (λ_i, v_i):

$$\Sigma v_i = \lambda_i v_i, \qquad z = V^{\top}(x - \bar{x}), \qquad (7)$$

where $V = [v_1, \dots, v_r]$ collects the top r eigenvectors, x is an input vector, $\bar{x}$ its mean, and z is the r-dimensional representation.

Summary.

TFT provides a flexible temporal modeling backbone, while ISF-MID (with mRMR and PCA as comparators) yields compact, informative inputs by mitigating redundancy and emphasizing relevant signals. This methodological foundation motivates the dataset construction and experimental design in the next sections.

3.3. Methodological contributions

This study follows a methodological design with several key contributions that improve the predictive performance and interpretability:

- **ISF-MID**: The Initial Selected Features–Mutual Information Difference (ISF-MID) as a redundancy-sensitive feature selection strategy adapted to the nature of financial time series data.
- **TFT Integration**: To make the best use of temporal and static covariates, ISF-MID is integrated into a TFT-based forecasting framework.
- **Comparative Benchmarking**: The introduced framework is systematically benchmarked with a range of classical ML (e.g., SVR, RF, DT) and advanced DL (e.g., LSTM, DNN)-based models along with alternative preprocessing methods (mRMR, PCA).
- **Assessment through Multiple Metrics**: A diverse range of evaluation metrics is used, from MAE, MSE and RMSE to MAPE, R^2, DA and SR, enabling a thorough examination of the model performance.
- **Improved Interpretability**: With integrated variable selection, dimensionality reduction, and attentional mechanisms in place, our framework offers transparent insights regarding the comparative significance of input features that influence performance as well as dynamic

transformations learned from market interactions.

3.4. Dataset specification

3.4.1. Dataset description

The empirical analysis uses 1,108 consecutive trading days of daily observations from the TSE. The dataset includes technical, market, and macroeconomic indicators that jointly characterize the Iranian equity market. The predictive task is to estimate the next-day return of the Tehran Exchange Dividend and Price Index (TEDPIX), serving as the target variable.

3.4.2. Feature overview

To emphasize the complementary roles of predictor variables in return prediction, we classify predictors into three categories:

- ***Market Indices***: Historical Returns of TEDPIX, the Price Index of 50 More Active Companies (PIC50), and the 30 Largest Companies Index (LCI30), which would represent market-wide momentum and capture sectoral trends.
- ***Technical Indicators***: Trend and volume indicators such as Trading Volume (TV), Moving Averages (MA), Exponential Moving Averages (EMA), Relative Difference in Percentage (RDP) and others, which are extensively used in short-horizon forecasting.
- ***Macroeconomic Variables***: External drivers, including OPEC oil prices (OP), USD/IRR (DP) and EUR/IRR (EP) exchange rates, and gold coin prices (GP), reflecting macro and commodity influences.

3.4.3. Tools and preprocessing environment

Feature construction follows Zhong and Enke (2017), which is implemented in R (4.3.2) and Excel. Preprocessing and analysis were performed in MATLAB (R2024b) and Python (3.12.1) (Table 1).

This diverse set of features provides a comprehensive representation of internal market dynamics and external shocks, which also introduces redundancy and collinearity, motivating the feature reduction strategies (ISF-MID, mRMR, PCA) described earlier.

3.4.4. Data preparation

3.4.4.1. Normalization

Continuous variables are rescaled using Min–Max normalization to $[0,1]$:

$$v_i^{(\mathrm{norm})} = \frac{v_i - \min(V)}{\max(V) - \min(V)}. \quad (8)$$

Definitions: v_i is the i-th observation of a given variable V; min(V) and max(V) denote the minimum and maximum over the training set for V, which prevents variables with large magnitudes (e.g., exchange rates) from dominating others (e.g., relative indices).

Table 1. Features included in the dataset with descriptions and calculation methods

Feature	Description	Calculation Method
Tehran Exchange Dividend and Price Index (TEDPIX)	Return on the current day	(p(t)-p(t-1))/p(t-1)
	Return on day t-1	(p(t-1)-p(t-2))/p(t-2)
	Return on day t-2	(p(t-2)-p(t-3))/p(t-3)
	Return on day t-3	(p(t-3)-p(t-4))/p(t-4)
Trading Volume (TV)	Relative change on day t-1	(p(t-1)-p(t-2))/p(t-2)
	Relative change on day t-2	(p(t-2)-p(t-3))/p(t-3)
	Relative change on day t-3	(p(t-3)-p(t-4))/p(t-4)
Index of 50 More Active Companies (PIC50)	Return on day t-1	(p(t-1)-p(t-2))/p(t-2)
	Return on day t-2	(p(t-2)-p(t-3))/p(t-3)
	Return on day t-3	(p(t-3)-p(t-4))/p(t-4)
Index of 30 large companies (LCI30)	Return on day t-1	(p(t-1)-p(t-2))/p(t-2)
	Return on day t-2	(p(t-2)-p(t-3))/p(t-3)
	Return on day t-3	(p(t-3)-p(t-4))/p(t-4)
Exponential Moving Average (EMA)	Four distinct EMAs of closing prices: 10, 20, 50, and 200 days	p(t)×(2/(10+1))+EMA10(t-1)×(1-2/(10+1))
		p(t)×(2/(20+1))+EMA20(t-1)×(1-2/(20+1))
		p(t)×(2/(50+1))+EMA50(t-1)×(1-2/(50+1))
		p(t)×(2/(200+1))+EMA200(t-1)×(1-2/(200+1))
Moving Average (MA)	Three distinct MAs: 5, 10, and 30 days	$MA_{t-1} = \frac{\sum_1^n p_{t-1}}{n}$
Relative Difference Percentage (RDP)	5-day relative difference (%)	(p(t)-p(t-5))/p(t-5)×100
	10-day relative difference (%)	(p(t)-p(t-10))/p(t-10)×100
	15-day relative difference (%)	(p(t)-p(t-15))/p(t-15)×100
	20-day relative difference (%)	(p(t)-p(t-20))/p(t-20)×100
Support/Resistance Index (MPP)	levels at 30 and 120 days	$MPP = \frac{p - p_{min}}{p_{max} - p_{min}}$
OPEC Oil Prices (OP)	Price on the current day and the past three days and return on day t-1	(p(t-1)-p(t-2))/p(t-2)
	Return on day t-2	(p(t-2)-p(t-3))/p(t-3)
	Return on day t-3	(p(t-3)-p(t-4))/p(t-4)
USD/IRR Exchange Rate (DP)	Price on the current day and the past three days and return on day t-1	P(t), P(t-1), P(t-2), P(t-3) (p(t-1)-p(t-2))/p(t-2)
	Return on day t-2	(p(t-2)-p(t-3))/p(t-3)
	Return on day t-3	(p(t-3)-p(t-4))/p(t-4)
EUR/IRR Exchange Rate (EP)	Price on the current day and the past three days and return on day t-1	P(t), P(t-1), P(t-2), P(t-3) (p(t-1)-p(t-2))/p(t-2)
	Return on day t-2	(p(t-2)-p(t-3))/p(t-3)
	Return on day t-3	(p(t-3)-p(t-4))/p(t-4)
Gold Coin Price (GP)	Price on the current day and the past three days and return on day t-1	P(t), P(t-1), P(t-2), P(t-3) (p(t-1)-p(t-2))/p(t-2)
	Return on day t-2	(p(t-2)-p(t-3))/p(t-3)
	Return on day t-3	(p(t-3)-p(t-4))/p(t-4)

3.4.4.2. Temporal windowing

To model sequential dependencies, the input samples are reshaped into sliding windows of length $L = 5$. Let d be the number of predictors and N the number of days; the resulting tensor has shape $(N - L + 1) \times L \times d$, enabling the TFT to capture short-term market memory.

3.4.4.3. Handling missing data

Missing values are imputed by forward fill, preserving temporal continuity without introducing synthetic noise, appropriate in financial time series, where the most recent value often carries a predictive signal.

3.4.4.4. Rolling window splitting

To mimic realistic trading conditions, a rolling (forward-expanding) split is used. Let $X_t \in \mathbb{R}^d$ denote the predictor vector at day t and $y_t \in \mathbb{R}$ the target return variable. For a split index t and horizon K,

$$\text{Training: } \{(X_1, y_1), \dots, (X_t, y_t)\}, \text{Testing: } \{(X_{t+1}, y_{t+1}), \dots, (X_{t+K}, y_{t+K})\}.$$

Evaluation is thus conducted on unseen, forward-looking data, avoiding information leakage under non-stationarity.

3.5. Methodology

This section presents the methodological design of the study, beginning with the proposed forecasting framework, followed by the baseline models, the hyperparameter tuning strategy, and finally the evaluation metrics. The structure ensures transparency, reproducibility, and alignment with best practices in financial forecasting.

3.5.1. Proposed framework

The proposed forecasting framework integrates ISF-MID and TFT. ISF-MID reduces the dimensionality of features by iteratively identifying predictors with strong mutual information with the target variable while penalizing redundancy. The selected features are then processed by the TFT, which models both short- and long-term dependencies and integrates static and dynamic covariates through attention and gating mechanisms. This hybrid design specifically addresses two challenges in financial forecasting: (i) noisy and high-dimensional predictors; and (ii) nonlinear temporal dependencies.

3.5.2. Baseline models for comparison

To validate the effectiveness of the ISF-MID+TFT framework, several ML and DL models were implemented as baselines:

- *MLP:* basic non-linear neural model;
- *RBF Network:* kernel-based approximation of local dynamics;
- *SVR:* robust kernel regression model;
- *DT and RF:* interpretable and ensemble tree-based learners;
- *DNN:* multi-layer feedforward architecture;
- *LSTM:* recurrent network for temporal dependencies.

This comparison ensures that TFT is evaluated against both classical ML and strong DL baselines, as summarized later in Table 2.

3.5.3. Model implementation and hyperparameter tuning

Hyperparameter optimization was performed using Optuna (Akiba et al., 2019), configured with the Covariance Matrix Adaptation Evolution Strategy (CMA-ES) sampler [20], an evolutionary optimization algorithm. The objective function minimized RMSE on the validation set over 100 trials. The hyperparameter search space was defined as follows:

- Learning rate: $[10^{-5}, 10^{-2}]$;
- Batch size: $\{32, 64, 128\}$;
- Hidden dimension: $[64, 256]$;

- Dropout rate: [0.0,0.5];
- Attention heads (TFT): {2,4,8}.

To ensure reproducibility, experiments were executed with fixed random seeds for both data shuffling and model initialization. Early stopping was applied with a patience of 15 epochs to prevent overfitting. The search converged after approximately 70 trials, with marginal gains thereafter, confirming that 100 trials provided a sufficient budget for reliable convergence.

The optimal configuration for the TFT was found to be as follows: learning rate = 0.001; batch size = 64; hidden dimension = 128; dropout = 0.1; and four attention heads. Comparable search spaces were defined for the baseline models. All experiments were implemented in Python (3.12.1) with PyTorch (2.2) on a workstation with an Intel i7 processor, 32GB RAM, and an NVIDIA GPU with 16GB memory.

3.5.4. Evaluation metrics

To ensure a rigorous and balanced assessment, we employed both statistical error measures (RMSE, MAE, MSE, MAPE, R^2) and domain-specific indicators (DA, SR), together with computational efficiency (training time). This combination captures accuracy, explanatory power, relevance of decisions, and practical feasibility. All error metrics are computed on Min–Max normalized returns, making them dimensionless and directly comparable across features.

The notation used in this section is as follows: $O_i \in \mathbb{R}$: observed value at index i; $P_i \in \mathbb{R}$: predicted value at index i; N: number of samples; $\bar{O} = 1/N \sum_{i=1}^{N} O_i$: mean of observed values; $\text{sign}(0)$: sign function; $\mathbb{1}\{0\}$: indicator function (1 if condition holds, 0 otherwise); μ_p: mean portfolio return per period; r_f: risk-free rate per period; σ_p: standard deviation of portfolio returns; T_{epoch}: training time per epoch (seconds).

3.5.5. Root mean squared error (RMSE)

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2} \tag{9}$$

Capturing the average magnitude of errors quadratically penalizes large deviations. It is useful when big mistakes are especially costly (Mohebbi et al., 2024).

3.5.6. Mean Absolute Error (MAE)

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}|O_i - P_i| \tag{10}$$

Equation 10 represents the mean absolute deviation, which is more robust to outliers than RMSE and directly interpretable in normalized units (Mohebbi et al., 2024).

3.5.7. Mean squared error (MSE)

$$\text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(O_i - P_i)^2 \tag{11}$$

Large errors are highlighted in Equation 11 due to squaring, forming the basis of RMSE and many least-squares objectives (Mohebbi et al., 2024).

3.5.8. Mean absolute percentage error (MAPE)

$$\text{MAPE} = \frac{100}{N}\sum_{i=1}^{N}\left|\frac{O_i - P_i}{O_i}\right| \quad (12)$$

Equation 12 provides a scale-free percentage error for cross-series comparison, and its interpretability is advantageous, though instability may arise when $O_i \approx 0$ (Mohebbi et al., 2025(a)).

3.5.9. Coefficient of determination (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(O_i - P_i)^2}{\sum_{i=1}^{N}(O_i - \bar{O})^2} \quad (13)$$

The proportion of variance is indicated in observations explained by the model; values closer to 1 imply stronger explanatory power (Mohebbi et al., 2025(b)).

3.5.10. Training time per epoch

Training time per epoch reports computational efficiency as T_{epoch} (s/epoch), reflecting scalability and deployability for real-time or large-scale applications (Mohebbi et al., 2025(a)).

3.5.11. Directional accuracy (DA)

$$\text{DA} = \frac{1}{N-1}\sum_{i=2}^{N}\mathbb{1}\{\text{sign}(O_i - O_{i-1}) = \text{sign}(P_i - P_{i-1})\} \quad (14)$$

Mean Directional Accuracy (MDA) measures the proportion of correctly predicted movement directions, which is particularly important in financial contexts, where the direction of change often outweighs precise error minimization (Costantini et al., 2016).

3.5.12. Sharpe ratio (SR)

$$\text{SR} = \frac{\mu_p - r_f}{\sigma_p} \quad (15)$$

Sharpe ratio (SR) evaluates risk-adjusted returns by comparing excess portfolio gains to volatility. Widely used in finance, values above 1 generally indicate favorable risk–return trade-offs and practical profitability (Bieganowski and Ślepaczuk, 2025).

4. Experimental results

4.1. Notation and units

O_i: observed return at index i; P_i: predicted return at index i; N: number of test samples; $\bar{O}$: mean of observed returns; MAE, RMSE: reported in the same unit as the output variable (Tehran Stock Exchange Price Index, TEPIX); MSE: reported in squared TEPIX units; MAPE: reported as a fraction in [0,1]; R: unitless; DA: reported in percent; SR: unitless; Time: seconds per epoch (s/epoch).

4.2. Benchmarking and proposed model

To determine the most effective model for predicting daily returns of the TSE index, several ML and DL models were compared. These included MLP, RBF networks, SVR, DT, RF, DNN, LSTM, and TFT. All these models have been widely applied in the financial forecasting literature.

Figure 2 presents an initial performance comparison, where the TFT demonstrated the lowest prediction error in multiple sets of characteristics. LSTM ranked second, consistent with its established effectiveness for financial time series. However, the TFT's capability to model complex temporal dependencies and to leverage both static and dynamic covariates yields higher accuracy.

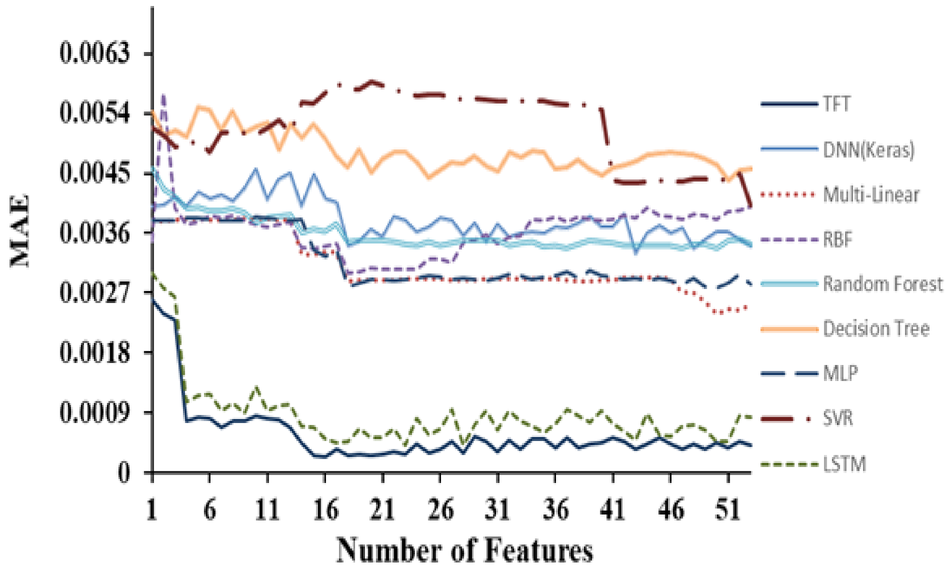

Figure 2. Comparative analysis of soft computing models for next-day stock index prediction

These findings validate the selection of the TFT as the proposed model for predicting the daily returns of the TSE index.

4.3. Performance comparison of TFT with other models

To evaluate its performance in predicting stocks, its performance is compared with cutting-edge deep and conventional machine learning algorithms, including LSTM, MLP, Random Forest, DNN, SVR, RBF, and DT. All of them have been compared with a range of performance metrics, which are MAE, MSE, RMSE, MAPE, R^2 Score, training time per epoch, DA, and SR. All are presented in a concise form in Table 2.

The results in Table 2 reveal that TFT achieves the best performance across all evaluation metrics, with LSTM as the closest competitor but still less efficient, while simpler models such as RF, MLP, and SVR show notably higher errors with minimum values for errors (MAE, RMSE) and a high value for R^2 (0.989), representing its high potential to track market trends. In addition, TFT outperforms in terms of DA (83.2%) and significantly outperforms models such as DT and SVR, which struggle to go beyond random chance. To avoid overestimation concerns, we note that DA is computed strictly on out-of-sample test windows, ensuring no look-ahead bias or data leakage. This provides a conservative but reliable measure of the accuracy of direction. Its value for SR (4.95) also confirms its performance in terms of risk-adjusted returns, making it more reliable for financial forecasting. Here, the SR was calculated under an idealized daily trading strategy (long/short at the close, risk-free rate set to zero, no transaction costs or slippage). Thus, the reported SR should be interpreted as an upper bound of achievable profitability; in practice, market frictions would slightly lower the ratio.

Table 2. Performance comparison of different models based on various evaluation metrics (errors reported in actual TEPIX units; MSE in squared units; MAPE as a fraction; DA in percent; Time in s/epoch).

Model	MAE	RMSE	MAPE	R²	DA (%)	SR	Time (s/epoch)
TFT	0.000	0.001	0.021	0.989	83.200	4.950	120
LSTM	0.000	0.001	0.029	0.976	78.500	3.610	110
MLP	0.002	0.003	0.052	0.945	72.300	2.350	70
RF	0.002	0.004	0.058	0.938	69.800	1.250	30
DNN	0.003	0.004	0.065	0.926	68.200	1.850	90
RBF	0.003	0.005	0.071	0.918	66.700	0.670	25
SVR	0.003	0.005	0.073	0.912	65.900	0.610	20
DT	0.004	0.006	0.082	0.895	63.400	0.400	5

Although the TFT model has the longest computation time per epoch (120s), its performance in all key evaluation metrics is worth its long training period. Computational intensity comes with the use of a self-attention mechanism and sequential feature extraction, but it can also detect complex trends in the market more effectively than traditional approaches. DT (5 s/epoch) and SVR (20 s/epoch) are more efficient computationally, but their predictive accuracy is lower than that of TFT. Ultimately, the trade-off between computation time and predictive power favors TFT, as its substantial accuracy gains outweigh the additional training cost.

4.4. Feature compact analysis: selection and reduction

Feature selection and reduction are significant steps in enhancing the performance of predictive models. In the case of this research work, two advanced statistical methods, mRMR and PCA, were used to choose the most relevant features of the data.

4.4.1. Feature selection using mRMR

As introduced in Section 3, mRMR was adopted for feature selection to maximize relevance to the target variable and minimize redundancy between feature inputs. Due to the high dimensions of financial information, mRMR in its traditional form was computationally expensive; for that reason, two approximation techniques, Mutual Information Difference (MID) and F-Test Correlation Difference (FCD), have been adopted for efficiency gains (Mohebi et al., 2022a).

For discrete variables, the mRMR criterion is approximated by using MID (Mohebi et al., 2022a):

$$MID(x_i) = I(x_i; y) - \frac{1}{|S|} \sum_{x_j \in S} I\left(x_i; x_j\right) \tag{16}$$

where:

- $I(x_i; y)$ is the mutual information between the feature x_i and the target variable y, measuring how much knowledge x_i reduces the uncertainty about y.
- $I(x_i; x_j)$ is the mutual information between a feature x_i and another selected feature x_j, capturing redundancy.
- $|S|$ is the number of selected features in the subset S.

Mutual information $I(A; B)$ is defined as (Mohebi et al., 2022b):

$$I(A; B) = \sum_{a \in A} \sum_{b \in B} p\,(a, b) \log \frac{p(a, b)}{p(a)p(b)} \tag{17}$$

where $p(a,b)$ is the joint probability distribution of A and B, while $p(a)$ and $p(b)$ are their respective marginal probabilities.

For continuous variables, mutual information is replaced by an F-test score, leading to the FCD approximation (Mohebbi Najm Abad and Soleimani, 2018):

$$FCD(x_i) = F(x_i, y) - \frac{1}{|S|}\sum_{x_j \in S} F\left(x_i, x_j\right) \tag{18}$$

where:

- $F(x_i, y)$ is the F-statistic, which measures the correlation between the feature x_i and the target variable y.
- $F(x_i, x_j)$ is the F-statistic that measures the correlation between a feature x_i and another selected feature x_j, indicating redundancy.
- $|S|$ is the number of selected features in the subset S.

The features were prioritized by the MID and FCD methods based on their relevance to the target variable and mutual independence. Prioritizations were preserved in the feature vector ζ. The subsets of characteristics ($1 \leq j \leq |\zeta|$) were evaluated using the TFT model to determine the optimal number of characteristics to forecast. The MAE was used as the performance metric.

The results, as illustrated in Figure 3, show that the accuracy of the model is initially enhanced by adding more selected features. However, after the 16th feature, the accuracy drops due to the addition of redundant or less significant features. Between the two mRMR approximation methods, MID outperformed FCD up to the 16th feature. Consequently, MID was selected as the preferred method for feature selection.

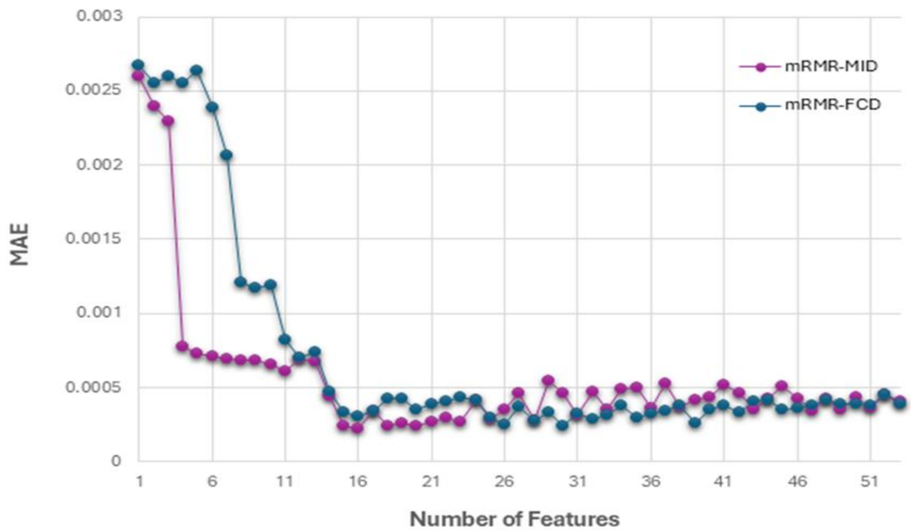

Figure 3. Impact of feature selection on prediction accuracy

The names of the first 16 best priority features using MID, which result in the highest accuracy in the prediction model, are presented in Table 3.

Table 3. Features prioritization based on importance scores

Rank	Feature Name	Description
1	RDP5	Relative Percentage Difference of Index Returns (5 Days)

2	U3	Euro Price Change (3 Days)
3	OP1	Crude Oil Price Change (1 Day)
4	PICS01	1-Day Return of Top 50 Active Stock Index
5	TV2	Relative Volume Change (2 Days)
6	RDP15	Relative Percentage Difference of Index Returns (15 Days)
7	D2	USD Price Change (2 Days)
8	D1	USD Price Change (1 Day)
9	D3	USD Price Change (3 Days)
10	OP3	Crude Oil Price Change (3 Days)
11	TV1	Relative Volume Change (1 Day)
12	RDP10	Relative Percentage Difference of Index Returns (10 Days)
13	GPd2	Gold Coin Price Change (2 Days)
14	TEDPIX3	3-Day Return of Overall Stock Index
15	OP2	Crude Oil Price Change (2 Days)
16	TEDPIX1	1-Day Return of Overall Stock Index

4.4.2. Feature extraction using PCA

PCA was used as a feature extraction algorithm to transform correlated input variables into an orthogonal principal component with preserved variance. The number of dimensions of the PCA was systematically varied to determine the optimal number of components, and the predictive performance of the TFT model was evaluated for each setting.

The results, as plotted in Figure 4, show that the accuracy of the model improves with additional PCA components to 30 dimensions. Beyond this, further increases in dimensionality yield diminishing returns in predictive capability. This suggests that PCA effectively reduces dimensionality while preserving significant information up to a certain threshold.

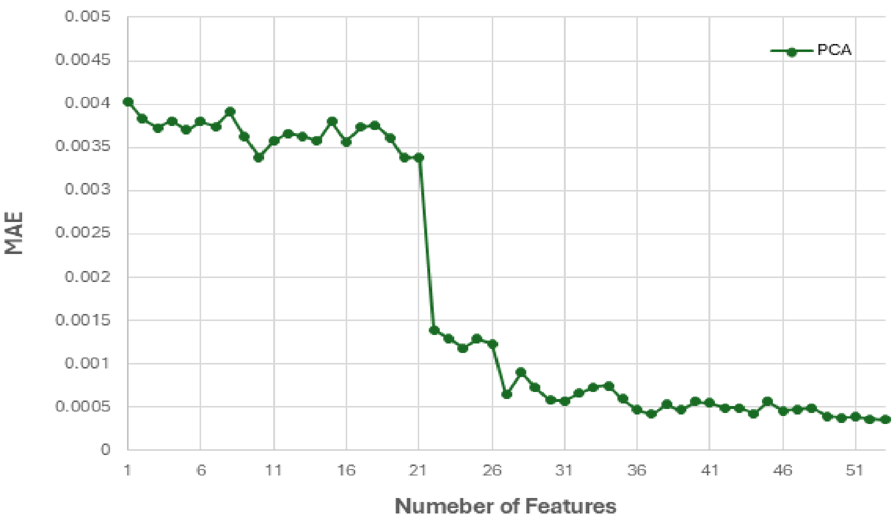

Figure 4. Impact of PCA Dimensionality on TFT Model Performance

4.4.3. Comparative analysis of mRMR and PCA

To compare the effectiveness of feature selection (mRMR) and feature extraction (PCA) techniques, their respective impacts on the predictive capacity of the TFT model were analyzed. The results of this comparison are shown in Figure 5.

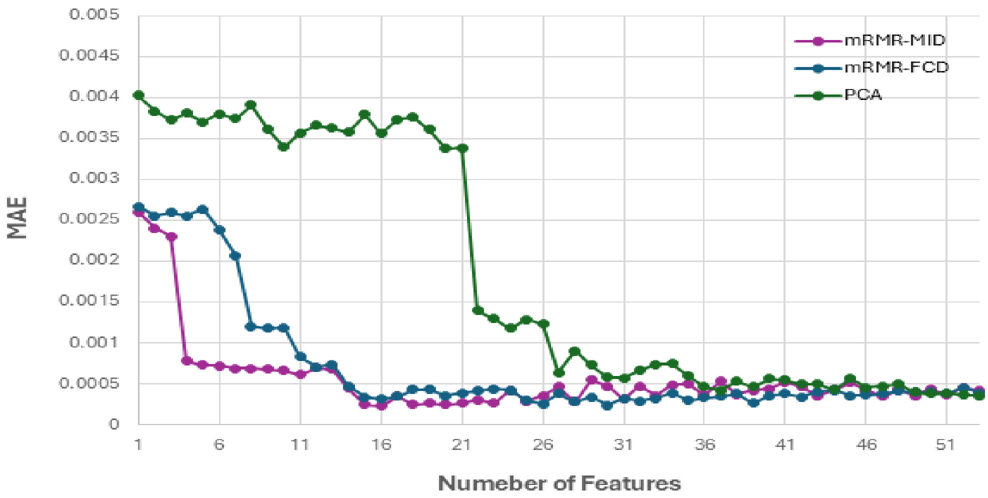

Figure 5. Comparison of mRMR and PCA techniques on TFT model performance

The findings indicate that while both techniques enhance model accuracy compared to using raw features, the mRMR-based feature selection method consistently outperforms PCA in terms of predictive accuracy. This advantage comes from the ability of mRMR to select the most relevant features explicitly. In contrast, PCA transforms the features into new components that may not necessarily capture the most predictive aspects of the data for the target variable. Based on this analysis, mRMR with the MID approximation was chosen as the preferred preprocessing technique for this study.

4.5. Impact of input feature selection on TFT model performance

To evaluate the contribution of individual input features to the precision of the prediction in the TFT model, three scenarios with a uniform and consistent network configuration were considered.

- ***Effective Features**:* In this scenario, only the effective features that have been proven to be significant for the prediction of the stock index were used as input to the model. In total, 25 such effective features have been considered.
- ***Effective Features and Statistical Features**:* Building on the same scenario, in this model, additional statistical characteristics derived from values of the historical parameters were included. There were additional features with information of historical perspective on parameter variations, increasing the total number of input features to 53.
- ***Selected Features with mRMR-MID**:* In the last scenario, to optimize the set of features, a subset of the most important features was derived using the mRMR-MID algorithm. A comparison of results showed that increasing the number of features to 16 increased the prediction accuracy, but no contribution to the accuracy was noticed for feature values over 16. Hence, the selection of 16 features was identified as the most effective approach to accurately predict the stock index. A comparison of these three scenarios is presented in Table 4.

Table 4. Results of the TFT model evaluation using different feature sets

Model Input Features	Number of Features	MAE	MSE	RMSE
Selected Influential Features	25	0.003	0.000	0.005
Influential Statistical Features	53	0.000	0.000	0.001
Optimally Selected Features	16	0.000	0.000	0.000

The results illustrate that the incorporation of statistical features significantly improves the precision of the prediction. Regardless of its success in representing long-term dependencies through self-attention mechanisms, including historical values of the stock index and other parameters with current values, leads to more reliable predictions. Furthermore, according to the error values in Table 4, selecting the 16 most relevant and highest-priority features, as determined by the mRMR-MID, generated a minimum model error.

4.6. Proposed algorithm for feature selection

In the proposed feature selection algorithm, an initial selection of feature sets was derived through a manual selection by the researcher. By carefully examining the variation in the pattern errors in Figure 5, an improvement in the accuracy of the prediction was observed with a feature count that increased to 16, and no further improvement was observed with any subsequent addition of features. Among the 16 initial features, the five most significant played the most important role in reducing model errors, which are:

- One-day lagged return of the top 50 active stocks index (PIC501)
- Two-day lagged USD price change (D2)
- Three-day lagged return of the overall index (TEDPIX3)
- Two-day lagged oil price change (OP2)
- One-day lagged return of the overall index (TEDPIX1)

Based on these observations, these five features were first selected, and then prioritization of the remaining features with the mRMR-MID algorithm was conducted. After that, these selected factors have been sequentially fed into the TFT model to predict the daily stock index return in Tehran. The flowchart of the proposed algorithm is presented in Figure 6.

In the implementation of ISF-MID, the redundancy penalty parameter α was empirically set to 0.5 as a balanced trade-off between the relevance of the features and the redundancy. Sensitivity tests with $\alpha \in \{0.3, 0.5, 0.7\}$ showed stable results, confirming robustness. To ensure comparability between features, all inputs were pre-processed using Min-Max normalization to $[0,1]$, as detailed in Section 4.2. This prevented scale disparities from biasing the mutual information estimates.

4.7. Comparative analysis for optimal TFT feature set

The following performance confirms the effectiveness of a larger number of prioritized features in improving the accuracy of the TFT model, which can be observed in Figure 7.a. The feature prioritization extracted via the proposed algorithm, known as the ISF-MID, is also displayed. Initially, the first five of them did not appear to be sufficient to provide a significant improvement in the accuracy of the prediction. However, when the sixth feature was incorporated, the necessity of the first five features was realized, resulting in a substantial improvement in the accuracy of the model, with lower error than the best result achieved in the previous phase.

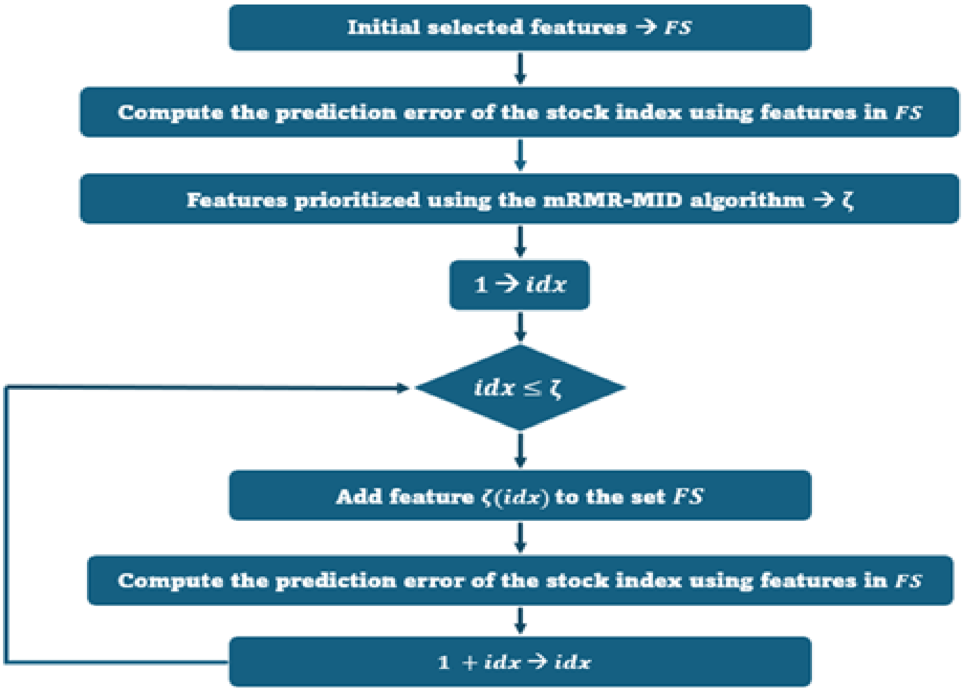

Figure 6. Flowchart of the proposed feature selection algorithm

This approach was further extended by introducing additional features to refine the stock index prediction model. As can be seen in Figure 7.a, both the sixth and eleventh characteristics, the difference in the five-day return in stocks and the three-day lagged oil price change, proved useful in reducing model error. Consequently, these two additional features were added to the initial selection, bringing the total to seven. The remaining were ranked with the mRMR-MID algorithm and incorporated into the TFT model. The evaluation result from this stage is shown in Figure 7.b.

As demonstrated in Figure 7.b, increasing the number of initial features to eight resulted in a substantial reduction in error, beyond which no further improvements were observed. Hence, the eighth feature, the relative change in the trading volume the previous day, was included in the selection at the last stage. Since no additional improvement in prediction accuracy could be noticed with an additional feature, these eight features were determined as the most effective features to predict the TSE index using the TFT model. The following is the ultimate selection of the chosen features:

- One-day lagged return of the top 50 active stocks index
- Two-day lagged USD price change
- Three-day lagged return of the overall index
- Two-day lagged oil price change
- One-day lagged return of the overall index
- Relative difference in five-day index return
- Three-day lagged oil price change
- Relative change in trading volume over one day

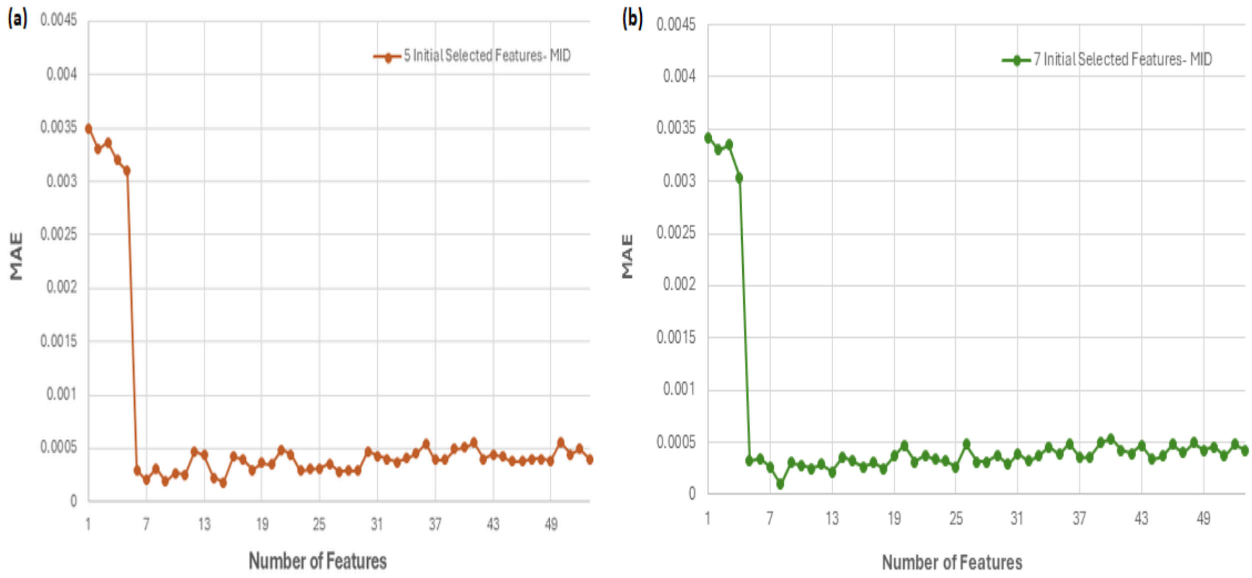

Figure 7. Evaluation of the TFT model with feature prioritization: (a) performance with five initial features (ISF-MID), (b) performance with seven initial features

A comparison was made with the mRMR-MID algorithm to assess the effectiveness of the proposed feature selection algorithm in reducing dimensions while improving prediction accuracy. Figure 8 illustrates the output of both feature selection algorithms. By comparing curves in Figure 8, a comparative analysis shows that the best performance of the TFT model with eight dimensions was achieved when the ISF-MID algorithm was utilized, confirming its effectiveness in feature selection. The prediction error values for feature selection cases for the prediction model with TFT are presented in Table 5. Experimental observations significantly improve predictive accuracy when using the proposed feature selection scheme. Specifically, when all 53 features are considered, an MAE value of 0.00042 is generated. However, when mRMR-MID feature selection is adopted, an improvement of 0.000239 is observed in the error value. Furthermore, the MAE decreased to 0.000230 by applying the proposed ISF-MID approach with a compact subset of six features. While this demonstrates the robustness of ISF-MID in extracting informative features, the six-feature configuration did not surpass the final eight-feature configuration. The best performance was ultimately achieved with ISF-MID using eight selected features, which produced the lowest MAE of 0.000097 and the highest R^2 (0.989). This establishes the eight-feature ISF-MID subset as the definitive optimal configuration for the TFT model in the study.

Table 5. Performance evaluation of the TFT model with different feature selection methods

Feature Selection Approach	Number of Features	MAE
All Features	53	0.000
mRMR-MID Selected Features	16	0.000
Proposed ISF-MID (6 Features)	6	0.000
Proposed ISF-MID (8 Features)	8	0.000

4.8. Evaluating training stability and prediction accuracy

A critical objective in model training is to have an optimal balance between underfitting and overfitting. The underfit and overfit balance is measured through observation of training and validation loss curves, and such plots illustrate the generalizability of a model to new observations and unseen data. A well-trained model has a decreasing trend in both losses up to a point where they stabilize, meaning convergence. Moreover, a minimal gap between training and validation loss implies that the model has effectively captured trends in the data without overfitting, as shown in Figure 9.

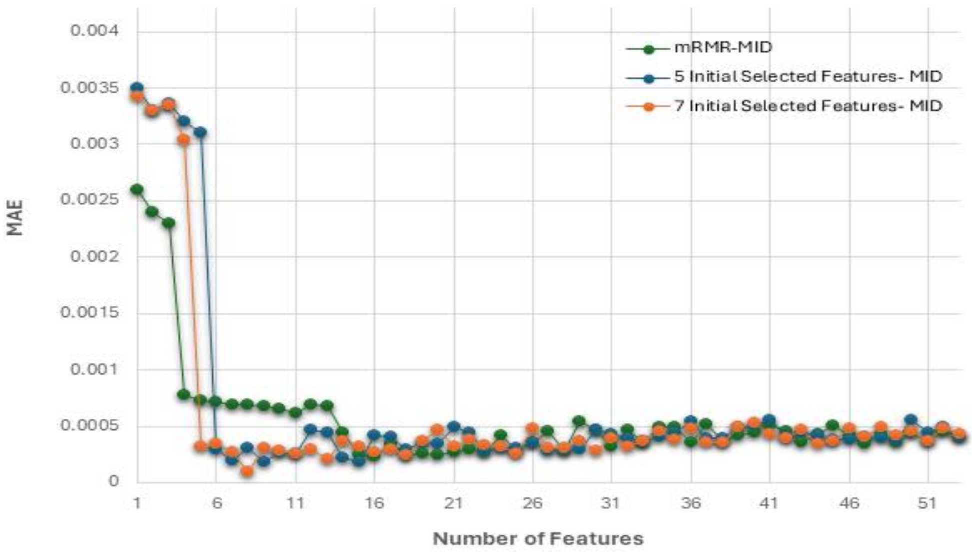

Figure 8. Comparative analysis of feature selection algorithms: ISF-MID vs. mRMR-MID in the TFT model

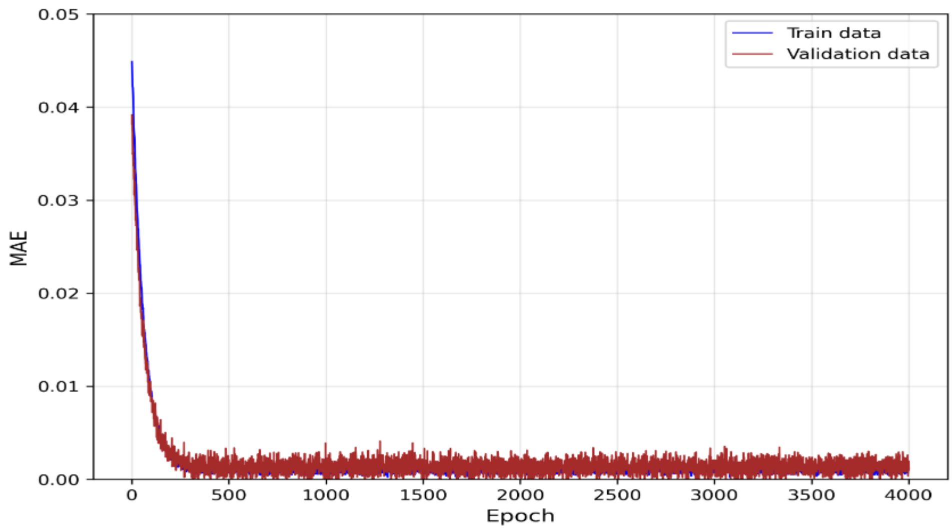

Figure 9. Training and validation loss curves for assessing model generalizability

Beyond loss curves, the predictive performance of the model should also be evaluated by comparing actual and predicted values for the stock index. To achieve this, the predicted values for the TFT model were compared with actual observations in several test samples. Figure 10.a shows a regression plot of test samples, illustrating how closely the model output aligns with the actual trends in the marketplace. Figure 10.b provides a comprehensive performance analysis in all test samples, further validating the accuracy of the model. The results show that with the optimally chosen features, the TFT model is most effective at predicting the real values of the stock market indices. This is corroborated by the minimal deviations within the prediction, which highlights the effectiveness of the model in describing market fluctuations. That means that the TFT model could be considered a powerful and reliable tool for financial time series prediction, as it manages to balance both training stability and predictive accuracy.

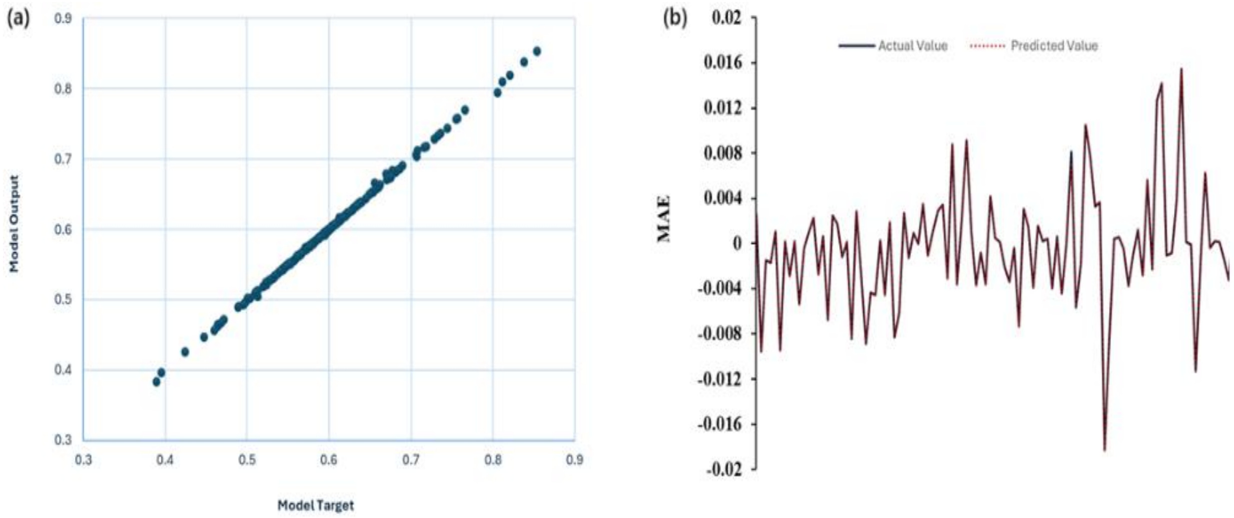

Figure 10. Evaluation of TFT model predictions: (a) Regression plot of predicted vs. actual values, (b) performance analysis across the full test set.

4.9. Statistical evaluation of the proposed model

The Friedman test is used to assess the significance of the results obtained from the predictions. It is a statistical method applied for two-way analysis of variance (for non-parametric data) using a ranking approach. In addition, the Friedman test can be used to compare the mean rankings of different groups. Table 6 presents the descriptive statistics of the errors in the models examined under the rolling window approach. As observed, the mean error in the TFT and LSTM models is lower than that of the other models, respectively. The Friedman test is a nonparametric test, equivalent to repeated measures analysis of variance (ANOVA), which is used to compare mean ranks between k variables (groups). In this test, a significance level less than 0.05 indicates a statistically significant difference in the errors of the models studied.

The results of the Friedman test are presented in Table 7. The findings indicate that the TFT model has a lower error compared to the other models examined. Since the significance level is below 0.05 and based on the mean ranks obtained, the TFT and LSTM models have the lowest mean ranks, respectively.

Table 6. Descriptive statistics of model errors in the rolling window approach

Prediction Model	Mean	Standard Deviation	Minimum	Maximum	Kurtosis	Skewness
MLP	0.002	0.003	0.000	0.056	22.733	3.617
SVR	0.003	0.003	0.000	0.062	47.313	4.771
DNN	0.003	0.003	0.000	0.040	18.916	3.256
Linear Regression	0.002	0.003	0.000	0.044	69.103	6.389
DT	0.004	0.005	0.000	0.055	19.705	3.472
RF	0.003	0.004	0.000	0.050	32.382	4.185
LSTM	0.000	0.001	0.000	0.002	156.975	10.054
TFT	0.000	0.001	0.000	0.025	464.283	20.063

Table 7. Comparison of median error in studied models based on the Friedman test

Model	Median Error	
MLP	5.070	
SVR	5.690	
DNN	5.290	
Linear Regression	4.510	
DT	5.870	
RF	5.100	
LSTM	2.840	
TFT	1.640	
Friedman Test Results	Chi-Square Statistic	1730.029
	Degrees of Freedom	7.000
	Significance Level	0.000

5. Conclusion

This study proposed a forecasting framework for daily returns of the TSE by combining domain-informed feature selection with deep learning. Among the candidate models, the TFT achieved the strongest performance by capturing temporal dependencies and nonlinear patterns in TEPIX data. An extensive set of 53 candidate features was initially constructed from 12 technical, market, and macroeconomic indicators. To reduce redundancy and improve efficiency, the mRMR-MID method was applied, resulting in 16 relevant features. Building on this, we introduced the ISF-MID algorithm, which integrates expert-informed initialization with redundancy-aware selection. ISF-MID achieved its best performance with only 8 features, producing the lowest MAE (0.000) and the highest R^2 (0.989). These results established TFT combined with ISF-MID (8 features) as the optimal configuration, demonstrating that a compact and well-prioritized subset can deliver highly accurate and interpretable forecasts in complex financial environments. In general, the integration of ISF-MID with TFT not only improved predictive accuracy but also provided a reproducible benchmark for researchers and a practical guideline for practitioners seeking efficient and interpretable models in financial forecasting.

Future work could extend this framework by incorporating transaction costs, liquidity constraints, and market frictions for realistic profitability assessment; adapting ISF-MID to nonlinear dependency measures or hybrid optimization strategies; and applying the framework to multi-market and cross-asset settings to evaluate its generalizability.

Declarations:

On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Akiba, T., Sano, S., Yanase, T., Ohta, T. and Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining (pp. 2623-2631). New York, NY. https://doi.org/10.1145/3292500.3330701
2. Alam, K., Bhuiyan, M. H., Haque, I. U., Monir, M. F. and Ahmed, T. (2024). Enhancing stock market prediction: A robust LSTM-DNN model analysis on 26 real-life datasets. *IEEE Access*, 12, pp. 122757-122768. https://doi.org/10.1109/ACCESS.2024.3434524
3. Awad, A. L., Elkaffas, S. M. and Fakhr, M. W. (2023). Stock market prediction using deep reinforcement learning. *Applied System Innovation*, 6(6), p. 106. https://doi.org/10.3390/asi6060106
4. Bao, W., Yue, J. and Rao, Y. (2017). A deep learning framework for financial time series using stacked autoencoders and long-short term memory. *PloS One,* 12(7), A. e0180944.

https://doi.org/10.1371/journal.pone.0180944
5. Bieganowski, B. and Ślepaczuk, R. (2025). Supervised autoencoder MLP for financial time series forecasting. *Journal of Big Data*, 12(1), p. 207. https://doi.org/10.1186/s40537-025-01267-7
6. Costantini, M., Cuaresma, J. C. and Hlouskova, J. (2016). Forecasting errors, directional accuracy and profitability of currency trading: The case of EUR/USD exchange rate. *Journal of Forecasting,* 35(7), pp. 652-668. https://doi.org/10.1002/for.2398
7. Dashtaki, S. M., Chagahi, M. H., Bahadori, A., Moshiri, B., Piran, M. J. and Montazeri, A. (2025). HSIF: A transformer-based cross-attention framework for cryptocurrency trend forecasting via multimodal sentiment–market fusion. *IEEE Access*, 13, pp. 156600-156612. https://doi.org/10.1109/ACCESS.2025.3605522
8. Davoud Abadi, M. (2025). Forecasting financial market indices using a transformer-based model (Master's thesis). Allameh Tabataba'i University, Tehran, Iran (In Persian).
9. Dos Santos, J. P. G., Moutinho, V. M. F., Madaleno, M. and Teixeira, Z. M. D. S. S. (2025). Examining the effect of macroeconomic, institutional, and capital market drivers on infrastructure investment. *Economic Analysis and Policy*, 86, pp. 165-190. https://doi.org/10.1016/j.eap.2025.02.038
10. Emami Gohari, H., Dang, X. H., Shah, S. Y. and Zerfos, P. (2024). Modality-aware transformer for financial time series forecasting. In Proceedings of the 5th ACM International Conference on AI in Finance (pp. 677-685). Association for Computing Machinery. https://doi.org/10.1145/3677052.3698654
11. Gandhudi, M., Alphonse, P. J. A., Velayudham, V., Nagineni, L. and Gangadharan, G. R. (2024). Explainable causal variational autoencoders based equivariant graph neural networks for analyzing the consumer purchase behavior in E-commerce. *Engineering Applications of Artificial Intelligence*, 136, A. 108988. https://doi.org/10.1016/j.engappai.2024.108988
12. Guo, Z., Ye, W., Yang, J. and Zeng, Y. (2017). Financial index time series prediction based on bidirectional two-dimensional locality preserving projection. In Proceedings of the 2017 IEEE 2nd International Conference on Big Data Analysis (ICBDA) (pp. 934 –938). IEEE, New York. https://doi.org/10.1109/ICBDA.2017.8078775
13. Hartanto, S. and Gunawan, A. A. S. (2024). Temporal fusion transformers for enhanced multivariate time series forecasting of Indonesian stock prices. *International Journal of Advanced Computer Science and Applications (IJACSA),* 15(7), pp. 1-22. https://doi.org/10.14569/IJACSA.2024.0150713
14. https://arxiv.org/abs/2503.18096
15. Jaiwang, G. and Jeatrakul, P. (2018). Enhancing support vector machine model for stock trading using optimization techniques. In 2018 International Conference on Digital Arts, Media and Technology (ICDAMT) (pp. 23-28). Phayao, Thailand. https://doi.org/10.1109/ICDAMT.2018.8376489
16. John, A. and Latha, T. (2023). Stock market prediction based on deep hybrid RNN model and sentiment analysis. *Automatika*, 64(4), pp. 981-995. https://doi.org/10.1080/00051144.2023.2217602
17. Kumarappan, J., Rajasekar, E., Vairavasundaram, S., Kotecha, K. and Kulkarni, A. (2024). Federated learning enhanced MLP–LSTM modeling in an integrated deep learning pipeline for stock market prediction. *International Journal of Computational Intelligence Systems*, 17(1), p. 267. https://doi.org/10.1007/s44196-024-00680-9
18. Kumbure, M. M., Lohrmann, C., Luukka, P. and Porras, J. (2022). Machine learning

techniques and data for stock market forecasting: A literature review. *Expert Systems with Applications,* 197, A. 116659. https://doi.org/10.1016/j.eswa.2022.116659
19. Li, Y., Ni, P. and Chang, V. (2020). Application of deep reinforcement learning in stock trading strategies and stock forecasting. *Computing*, 102(6), pp. 1305-1322. https://doi.org/10.1007/s00607-019-00773-w
20. Liang, X., Ge, Z., Sun, L., He, M. and Chen, H. (2019). LSTM with wavelet transform based data preprocessing for stock price prediction. *Mathematical Problems in Engineering,* 2019(1), A. 1340174. https://doi.org/10.1155/2019/1340174
21. Lim, B., Arık, S. Ö., Loeff, N. and Pfister, T. (2021). Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International journal of forecasting*, 37(4), pp. 1748-1764. https://doi.org/10.1016/j.ijforecast.2021.03.012
22. Liwei, T., Li, F., Yu, S. and Yuankai, G. (2021). Forecast of LSTM-XGBoost in Stock Price Based on Bayesian Optimization. *Intelligent Automation & Soft Computing*, 29(3), pp. 855–868. https://doi.org/10.32604/iasc.2021.016805
23. Mohebbi Najm Abad, J. and Soleimani, A. (2018). Novel feature selection algorithm for thermal prediction model. *IEEE Transactions on Very Large-Scale Integration (VLSI) Systems,* 26(10), pp. 1831-1844. https://doi.org/10.1109/TVLSI.2018.2841318
24. Mohebbi, M. R., Kafash, E. and Döller, M. (2025a). Multi-Agent Trajectory Prediction for Urban Environments with UAV Data Using Enhanced Temporal Kolmogorov-Arnold Networks with Particle Swarm Optimization. In ICAART (2) (pp. 586-597). ScitePress, Setúbal, Portugal. https://doi.org/10.5220/0013243100003890
25. Mohebbi, M. R., Klinger, J., Abad, J. M. N., Döller, M. and Tavasoli, M. (2025b). Advanced driving behavior analysis through Kolmogorov-Arnold network and UAV traffic data. In International Conference on Machine Learning and Computing (pp. 305-322). Cham: Springer Nature Switzerland, Switzerland. https://doi.org/10.1007/978-3-031-94892-3_23
26. Mohebbi, M. R., Zada, M. J. H., Rostami-Shahrbabaki, M., Döller, M. and Yamnenko, I. (2024). Network traffic co-movement assessment via oriented basis signal processing and ensemble decision trees. In 2024, IEEE 27th International Conference on Intelligent Transportation Systems (ITSC) (pp. 1948-1955). IEEE, Piscataway, NJ. https://doi.org/10.1109/ITSC58415.2024.10919555
27. Mohebi, S., Fadaeinejad, M. E., Osoolian, M. and Hamidizadeh, M. R. (2022a). Feature Selection for the Prediction Model of the Tehran Stock Exchange Index by Dimensionality Reduction Techniques. *Financial Research Journal,* 24(4), pp. 577-601. https://doi.org/10.22059/frj.2021.325675.1007202
28. Mohebi, S., Fadaeinezhad, M. and Osoolian, M. (2022b). Predicting the Tehran Stock Exchange Index using support vector regression; Based on the dimension reduction technique. *Financial Management Strategy*, 10(3), pp. 1–26 (In Persian). https://doi.org/10.22051/jfm.2022.35543.2528
29. Nabipour, M., Nayyeri, P., Jabani, H., Mosavi, A., Salwana, E. (2020). Deep learning for stock market prediction. *Entropy*, 22(8), p. 840. https://doi.org/10.3390/e22080840
30. Niu, H., Xu, K. and Wang, W. (2020). A hybrid stock price index forecasting model based on variational mode decomposition and LSTM network: A hybrid stock price index forecasting model based on variational mode decomposition and LSTM network. *Applied Intelligence*, 50(12), pp. 4296-4309. https://doi.org/10.1007/s10489-020-01814-0
31. Pourroostaei Ardakani, S., Du, N., Lin, C., Yang, J. C., Bi, Z. and Chen, L. (2023). A federated learning-enabled predictive analysis to forecast stock market trends. *Journal of*

Ambient Intelligence and Humanized Computing, 14(4), pp. 4529-4535. https://doi.org/10.1007/s12652-023-04570-4
32. Sezer, O. B., Gudelek, M. U. and Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. *Applied Soft Computing*, 90(1), A. 106181. https://doi.org/10.1016/j.asoc.2020.106181
33. Sharma, R. and Mehta, K. (2024). Stock market predictions using deep learning: developments and future research directions. *Deep Learning Tools for Predicting Stock Market Movements*, 1(1), pp. 89-121. https://doi.org/10.1002/9781394214334.ch4
34. Sun, Y., Yuan, H. and Xu, F. (2025). Financial sentiment analysis for pre-trained language models incorporating dictionary knowledge and neutral features. *Natural Language Processing Journal*, 11(1), A. 100148. https://doi.org/10.1016/j.nlp.2025.100148
35. Tulcanaza Prieto, A. B. and Lee, Y. H. (2019). Determinants of stock market performance: VAR and VECM designs in Korea and Japan. *Global Business & Finance Review (GBFR),* 24(4), pp. 24-44. https://doi.org/10.17549/gbfr.2019.24.4.24
36. Wang, C., Chen, Y., Zhang, S. and Zhang, Q. (2022). Stock market index prediction using deep Transformer model. *Expert Systems with Applications*, 208, A. 118128. https://doi.org/10.1016/j.eswa.2022.118128
37. Wei, D. (2019). Prediction of stock price based on an LSTM neural network. In the 2019 International Conference on Artificial Intelligence and Advanced Manufacturing (AIAM) (pp. 544-547). IEEE, New York. https://doi.org/10.1109/AIAM48774.2019.00113
38. Wei, X., Chen, W. and Li, X. (2021). Exploring the financial indicators to improve the pattern recognition of economic data based on machine learning. *Neural Computing and Applications,* 33(2), pp. 723-737. https://doi.org/10.1007/s00521-020-05094-0
39. Wei, Y., Gu, X., Feng, Z., Li, Z. and Sun, M. (2024). Feature extraction and model optimization of deep learning in stock market prediction. *Journal of Computer Technology and Software*, 3(4), 1-22. https://doi.org/10.5281/zenodo.13622489
40. Wu, Y. (2023). Comparison between transformer, informer, autoformer and non-stationary transformer in financial market. *Applied and Computational Engineering*, 29, pp. 68-78. https://doi.org/10.54254/2755-2721/29/20230874
41. Wu, Y., Fu, Z., Liu, X. and Bing, Y. (2023). A hybrid stock market prediction model based on GNG and reinforcement learning. *Expert Systems with Applications*, 228, A. 120474. https://doi.org/10.1016/j.eswa.2023.120474
42. Yang, J., Li, P., Cui, Y., Han, X. and Zhou, M. (2025). Multi-sensor temporal fusion transformer for stock performance prediction: An adaptive Sharpe ratio approach. *Sensors*, 25(3), p. 976. https://doi.org/10.3390/s25030976
43. Zhang, J., Cui, S., Xu, Y., Li, Q. and Li, T. (2018). A novel data-driven stock price trend prediction system. *Expert Systems with Applications,* 97(1), pp. 60-69. https://doi.org/10.1016/j.eswa.2017.12.026
44. Zheng, H., Wu, J., Song, R. and Guo, L. (2024). Predicting financial enterprise stocks and economic data trends using machine learning time series analysis. *Applied and Computational Engineering*, 87(1), pp. 26–32. https://doi.org/10.20944/preprints202407.0895.v1
45. Zhong, X. and Enke, D. (2017). Forecasting daily stock market return using dimensionality reduction. *Expert Systems with Applications*, 67(2), pp. 126-139. https://doi.org/10.1016/j.eswa.2016.09.027

IJAAF; Vol. 10, No. 2, Spring 2026, Serial 35

Ferdowsi University of Mashhad

Iranian Journal of
Accounting, Auditing & Finance

Quarterly

RESEARCH ARTICLE

The Effect of Working Capital Management on Cost Stickiness Considering the Mediating Role of Financial Constraints

Mehdi Kazempour Barough, Hamzeh Didar*, Farzad Ghayour

Department of Accounting, Faculty of Economics and Management, Urmia University, Iran

How to cite this article:
Kazempour Barough, M. , Didar, H. and Ghayour, F. (2026). The Effect of Working Capital MaThe Effect of Working Capital Management on Cost Stickiness Considering the Mediating Role of Financial Constraintsanagement on Cost Stickiness Considering the Mediating Role of Financial Constraints. Iranian Journal of Accounting, Auditing and Finance, 10(2), 153-170. doi: 10.22067/ijaaf.2026.47673.1598
https://ijaaf.um.ac.ir/article_47673.html

ARTICLE INFO

Article History
Received: 2025-05-11
Accepted: 2025-10-07
Published online: 2026-04-12

Abstract

This study examined the effect of working capital management (WCM) policies on cost stickiness, providing the first direct empirical quantification of whether and how financial constraints mediate this relationship in a major emerging market. Building on established trade-off, agency, and adjustment-cost theories, we hypothesize that aggressive WCM primarily reduces cost stickiness for efficiency, but that financial constraints partially offset this gain. Using 1,848 firm-year observations from 132 Tehran Stock Exchange (TSE)–listed firms (2010–2023), we estimated multiple regression and a covariance-based SEM. More aggressive WCM is associated with lower cost stickiness. Critically, the mediation analysis indicates partial mediation; aggressive WCM is linked to higher financial constraints, and these constraints, in turn, amplify cost stickiness, thereby quantifying the trade-off between direct efficiency gains and indirect financial rigidities. Our findings establish an empirical foundation for advising managers to adopt aggressive WCM strategies only when coupled with strategies to mitigate financial constraint risk. Policymakers, particularly in SME contexts, should design targeted liquidity provision mechanisms to unlock the full cost-flexibility potential of WCM.

Keywords:
Cost Stickiness, Financial Constraints, Working Capital Management

https://doi.org/10.22067/ijaaf.2026.47673.1598

CC BY

NUMBER OF REFERENCES
45

NUMBER OF FIGURES
1

NUMBER OF TABLES
5

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

*Corresponding Author: Hamzeh Didar
Email: h.didar@urmia.ac.ir
Tel: 09143413984
ORCID: 0000-0003-4651-7729

1. Introduction

In corporate finance, WCM, or strategic management of a firm's short-term liabilities and assets, has grown to become a fundamental area of operational performance, liquidity, and value generation (Gitman, 1974; Deari et al., 2024). Effective WCM enables firms to meet short-term obligations by improving cash flow and profitability, primarily through reductions in the Cash Conversion Cycle (CCC) (Huynh et al., 2025). While historically viewed as a routine money-generating activity, WCM is nowadays a managerial option at the mercy of the risk-taking nature of the manager, industry rivalry, and the overall macro-financial environment (Filbeck and Krueger, 2005; Prasad et al., 2019).

Concurrently, a growing cost accounting literature has examined asymmetric cost behavior, particularly cost stickiness, the mechanism through which costs decrease less when activity declines than they increase when activity rises by the same margin (Anderson et al., 2003). These asymmetries challenge conventional proportional cost adjustment-based models and have significant managerial choice implications, earnings forecasting, and firm resilience (Banker and Byzalov, 2014; Deneckere and Kovenock, 1996; Ballas et al., 2022). Despite the large amount of research conducted on WCM and asymmetric cost behavior independently, relatively little is known regarding the interplay between WCM and asymmetric cost behavior.

Causes of cost stickiness are attributed to discretionary economic decisions resulting from resource flexibility for change costs of resources such as capacity, labor, and infrastructure (Anderson and Lanen, 2007). Theoretical models of cost stickiness (Anderson et al., 2003) focus on the fact that managers typically avoid cutting costs in a depression due to expected future recoveries or fear of frictions imparted by cost adjustment (Restuti et al., 2022). Simultaneously, WCM approaches ranging from aggressive (low cost usage, liquidity-constrained) to conservative (high liquidity, risk-sensitive) set up a company's ability and desire to modify its cost structure (Hung and Su Dinh, 2022). Prior studies (e.g., Berg et al., 2024) have shown that cost stickiness is less pronounced in aggressive WCM firms due to greater cost sensitivity to promote efficiency. Conversely, conservative WCM practices can solidify cost stickiness by strategic resource holding back (Mirón Sanguino et al., 2024). Since WCM decisions involve short-term resource allocation and liquidity control, they directly affect how quickly a firm can adjust its operating costs when activity levels change.

The relationship gets complex in the case of financially constrained firms. For financially constrained firms with limited access to external capital, particularly small companies, cost adjustments would occur faster, irrespective of risk management purposes, because such firms do not possess the ability to buffer economic shocks (Hadlock and Pierce, 2010).

Thus, the definition using financial constraints yields implications about direct influence on costs behavior and mediation mechanism, contending that these are mechanisms intervening between WCM strategies' effect and asymmetric cost behavior (a central yet unexplored process) (Ibrahim et al., 2022). This has received more attention in recent years, but no studies have tested the mediating effect of financial constraints on asymmetric cost behavior in terms of WCM strategies. Most empirical studies have considered the effects of WCM and financial constraint individually and sometimes side by side, but never in terms of how they might interact or condition one another (Jabbouri et al., 2024). This, indeed, is a very critical omission because aggressive WCM, aimed at cost reduction, may expose an enterprise to financial constraints, thus limiting cost reactions. Conversely, conservative WCM shields a firm from such constraints but, at the same time, may impose cost rigidity (Ujah et al., 2020). Without an understanding of how financial limitations mediate this interaction, managerial and policy implications stemming from existing literature may be suboptimal or even inaccurate.

This study bridges this gap by investigating whether financial limitations account for the interaction between WCM strategies and cost stickiness (Chen and Ma, 2021), which also considers whether aggressive WCM policies impact cost asymmetry indirectly through financing constraints (Jabbouri et al., 2024). Managerial cost behavior causal mechanisms are tested through a mediation analysis framework (Baron and Kenny, 1986; Hayes and Scharkow, 2013). To this effect, the study aims to determine how and why WCM affects cost adjustments under financial frictions, gaining insight into firm behavior under regimes of uncertainty.

Accordingly, this study seeks to answer the following research question: To what extent do financial constraints mediate the relationship between WCM strategies and cost stickiness? By addressing this question, the study contributes to the literature in several important ways. Specifically, it offers three distinct contributions:

First, it establishes an empirical foundation for the strategic WCM trade-off by quantifying, for the first time in an emerging market context, focusing on firms listed in the Tehran Stock Exchange, the precise extent to which constraint-induced rigidity offsets WCM efficiency gains.

Second, it theoretically reconciles the conflicting effects of aggressive WCM by demonstrating the mechanism through which the same policy simultaneously drives the efficiency channel and the vulnerability channel, thereby enriching the application of Adjustment Cost Theory.

Third, this research provides specific managerial guidance, advising firms to adopt WCM efficiency initiatives only in conjunction with financing strategies aimed at mitigating liquidity risk. This also lays the groundwork for policy-targeted interventions to support small and medium-sized enterprises (SMEs) facing financial constraints.

2. Literature review

2.1. Theoretical basis

The present study is based on three theoretical frameworks: the liquidity-profitability trade-off theory in WCM, the adjustment cost theory in cost behavior, and the agency theory. Liquidity-profitability trade-off theory: This theory states that firms face a fundamental conflict between maintaining liquidity and achieving profitability (Deloof, 2003; Baños-Caballero et al., 2014). In aggressive WCM policies, the level of investment in current assets (such as inventory and accounts receivable) is minimized while reliance on current liabilities (such as accounts payable) increases (Agegnew, 2019). Adopting such a strategy boosts free cash flow at the expense of liquidity risk associated with greater usage of external financing channels, which also worsens financial frictions (Atanasova and Willeboordse, 2025). Aggressive working capital policies improve efficiency and profitability ultimately, but from a trade-off perspective, they also inherently increase liquidity risk and dependence on external finance, increasing the firm's vulnerability to financial constraints. This provides the theoretical underpinnings for why it is important to explore financial constraints as a mediating mechanism.

By keeping a large amount of current assets, conservative policies might reduce profitability (Singh et al., 2017), but they can ensure more liquidity. The major theory explaining the cost stickiness is based on the fact that there are adjustment costs. When sales fall, managers are reluctant to take down resources such as skilled labor or key supplier relationships too quickly, for the repositioning costs (e.g. of hiring and training) in the future can be high (Anderson et al., 2003; Banker and Chen, 2006). Simultaneously, agency theory presents a complementary explanation and posits that managers might be motivated to hoard resources beyond the optimal level for self-serving reasons, e.g. establishing empires (empire building) or augmenting job security. Such behavior can create cost stickiness even when it does not enhance shareholder value (Chen et al., 2012; Jensen and Meckling, 1976), a characteristic that we continue in our model.

When facing financial constraints, managers may have a limited capacity to make strategic resource adjustments: costs of adjustment become more binding and agency considerations play an influence in terms of choosing whether to retain or release resources, potentially heightening cost stickiness even during periods of revenue declines.

2.2. The link between WCM and cost stickiness

The strategic approach to WCM directly affects a company's cost structure (Vlismas, 2024). According to the principles of the trade-off theory, the selection of a WCM strategy is not merely operational; rather, it reflects a fundamental balance between efficiency and financial vulnerability, which subsequently shapes managerial cost decisions. Firms adopting aggressive WCM policies are philosophically and operationally oriented toward efficiency. Managers in such firms typically possess strong capabilities in eliminating excess inventory, accelerating receivables collection, and managing short-term liabilities, which naturally extend to operating cost management (Frankel et al., 2017; Aktas et al., 2015).

Further, the causal link between WCM (particularly the cash conversion cycle) and cost stickiness can be explained through the adjustment cost theory, which posits that constrained liquidity increases the marginal cost of adjusting operational resources and influences managers' choices between reducing costs and preserving capacity (Jose et al., 1996; He et al., 2020). An efficiency-oriented managerial mindset, therefore, reduces the likelihood of accumulating slack resources and increases managers' willingness to accept adjustment costs to align the cost structure with declining revenues.

Empirical evidence supports these mechanisms in which firms with shorter cash conversion cycles, an indicator of aggressive WCM, tend to exhibit lower levels of cost stickiness (Berg et al., 2024). Combining insights from the trade-off theory, the adjustment cost theory, and the agency theory, we can articulate the theoretical rationale for linking aggressive WCM, financial constraints, and cost stickiness.

Aggressive WCM, typically reflected in a shorter cash conversion cycle (CCC), a lower net trade cycle (NTC), and a reduced working-capital-to-total-assets ratio (WCTA), modifies the firm's operating resource base by reducing cash tied up in inventory and receivables while increasing reliance on spontaneous financing through accounts payable (Baños-Caballero et al., 2014; Aktas et al., 2015). From the perspective of adjustment cost theory, reducing investment in working capital limits slack resources and prevents the buildup of idle capacity. Consequently, when sales decline, managers face fewer excess inputs and encounter lower adjustment costs associated with labor shedding, capacity downsizing, or supplier renegotiation (Anderson et al., 2003).

Under this resource-lean structure, the marginal cost of adjusting operating resources is lower, thereby weakening the conditions that typically generate asymmetric cost behavior. Therefore, in line with predictions of the aforementioned theories, firms following aggressive WCM policies are expected to exhibit less cost stickiness.

Combining these perspectives, it is argued that aggressive WCM, by reducing cash and inventory holdings (liquidity-profitability trade-off), reduces adjustment costs when resources are downsized (adjustment cost theory), while managerial agency considerations can also affect the excess of resources held (agency theory). These mechanisms collectively reduce cost stickiness and form the basis of hypothesis H1.

The first hypothesis, ensuing from these findings, is the following:

H1: Firms with more aggressive working capital management experience less cost stickiness.

2.3. The mediating role of financial constraints

Aggressive WCM operationalized by a shorter CCC, a lower NTC, and a lower WCT reduces the firm's holdings of cash, inventories and receivables, while increasing reliance on spontaneous and short-term liabilities (accounts payable) (Chow and Fung, 2000; Bottazzi et al, 2014). Empirically and theoretically, these three WCM proxies capture complementary aspects of the same policy choice. CCC and NTC reflect the timing of cash flows (days inventory and receivables relative to payables), whereas WCTA captures the intensity of current-asset investment relative to firm scale (Deloof, 2003; Baños-Caballero et al., 2014). By compressing internal liquid buffers and elevating dependence on external or supplier financing, aggressive WCM increases a firm's exposure to funding shortfalls and therefore the degree of financial constraint (Path A).

When financial constraints bind, firms face limited access to external funds needed to absorb the one-time, fixed costs of re-scaling operations (severance payments, contract termination costs, plant retooling, supplier renegotiation). Under the adjustment-cost theory, these fixed costs determine whether managers can economically implement immediate, large-scale adjustments; when internal and external liquidity are scarce, managers are unable to bear those upfront costs and therefore postpone or ration adjustments. The net consequence is that costs remain relatively inflexible in the face of declining sales, i.e., financial constraints amplify cost stickiness (Path B), because constrained firms cannot pay the fixed adjustment costs required to reduce committed inputs (Hadlock and Pierce, 2010).

Hence, full consideration of financial constraints would involve two relationships that need to be tested; aggressive WCM lowers liquidity and increases the constraint (Path A), whereas financial constraints limit managers' motivation to downsize in the face of declining sales (Path B), thereby strengthening cost stickiness.

Thus, the second hypothesis of this research reads as follows:

H2: Financial constraints mediate the relationship between aggressive working capital management and cost stickiness.

In summary, aggressive WCM (independent variable) affects cost stickiness (dependent variable) both directly and indirectly through financial constraints (mediator), forming the conceptual framework of this study in Figure 1.

3. Methodology

3.1. Sample selection

To achieve the objectives of this study, all firms listed on the TSE were considered as the population, with the sample selected systematically according to the following exclusion criteria:

1. Listed on the TSE from 2010 to 2023.
2. Fiscal years ending on Esfand 29 (March 19).
3. Perform as manufacturing companies, dropping off firms dealing with finance, including the banks, insurance companies, and investment firms.
4. Remaining listed on the TSE through 2023 with continuous operations.
5. Experiencing no transaction halts longer than six months.

Based on the above criteria, the final sample consisted of 132 manufacturing firms, yielding a total of 1,848 firm-year observations between 2010 and 2023. The Consumer Price Index (CPI), which tracks overall price fluctuations over time consistently, was used to deflate data for adjusting inflationary measures. Thus, the two measures of WCM can be consistent across sectors. In addition, we performed winsorization at the 1st and 99th percentiles to mitigate extreme values. Hypotheses were tested via multiple regression models and structural equation modeling (SEM) conducted in Stata software.

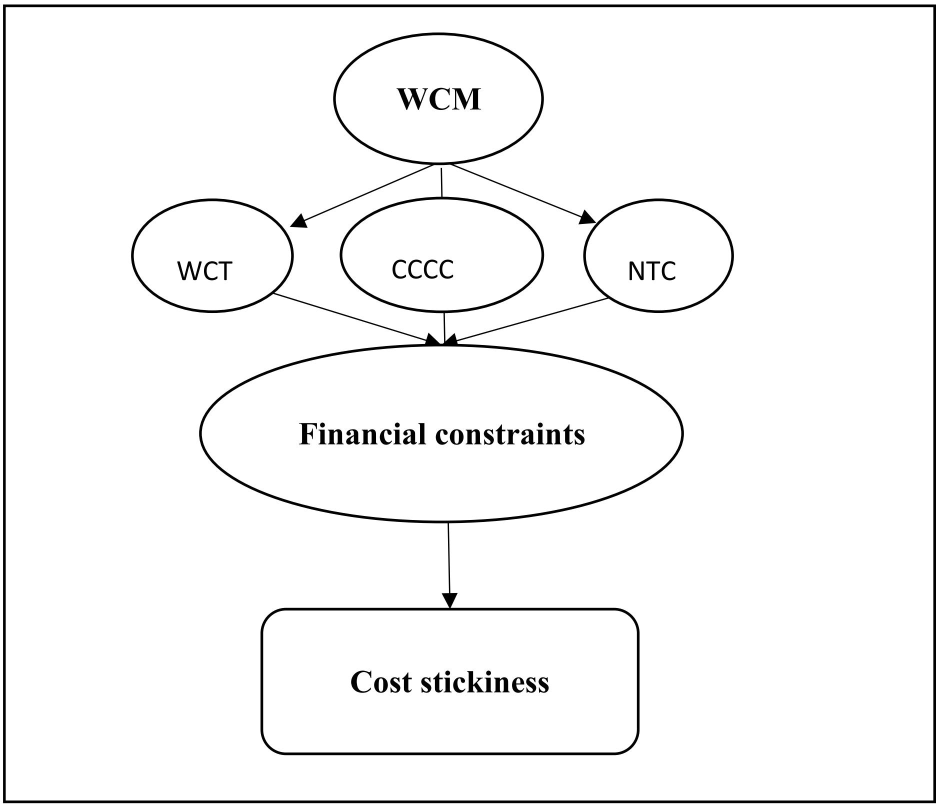

Figure 1. The conceptual framework

3.2. Research design

Model 1: Direct Effect of WCM on cost stickiness

In this study, an attempt was made to incorporate the model presented by Anderson et al. (2003) and, with reference to the study by Berg et al. (2024), to examine the effect of working capital management on asymmetric cost behavior in companies, which was the target of the first hypothesis. This model can be represented as follows:

$$\Delta\, ln\, COST_{i,t} = \beta_0 + \beta_1 \Delta\, ln\, SALES_{i,t} + \beta_2 D_{i,t} \times \Delta\, ln\, SALES_{i,t} + \beta_3 D_{i,t} \times \Delta\, ln\, SALES_{i,t} \times WCM_{i,t} + \beta_4 WCM_{i,t} + \beta_5 FCF_{i,t} + \beta_6 EMINT_{i,t} + \beta_7 SIZE_{i,t} + \beta_8 ROA_{i,t} + \beta_9 Cash\ holding_{i,t} + \beta_{10} \Delta GDP_{i,t} + \varepsilon_{i,t}$$

The regression equations are estimated three times by replacing WCM with CCC, NTC, and WCTA.

3.2.1 Variable definitions and measurement (Model 1)

3.2.1.1. Dependent variable

Change in Logarithmic Operating Costs (ΔlnCOST): This variable, known as the logarithmic change in operating costs, measures the change in a firm's operating costs from the previous year to the current year. This indicator was first used in the cost stickiness model by Anderson et al. (2003) and is used in most studies related to asymmetric cost behavior (Anderson et al, 2003).

$$\Delta\, ln\, COST = ln\, COST_{i,t} - ln\, COST_{i,t-1}$$

3.2.1.2. Independent variable

Working capital management (WCM): In order to conduct an analysis on the first hypothesis, the three main dimensions of WCM, namely, WCTA, CCC, and NTC, are examined. Based on the background of studies, CCC, as a common indicator used in WCM studies (Liu et al., 2024), focuses on the time and liquidity cycle, and WCTA, as a proxy for the liquidity health of the company, provides a comprehensive view of the current financial structure, and NTC, which is a relative measure of working capital, is useful for studying sales and comparability (Berg et al., 2024).

WCM: calculated using three metrics: WCTA, CCC, and NTC.

Working capital to total assets ratio (WCTA): WCTA is a widely used liquidity indicator that measures the proportion of a firm's total assets financed through net working capital, which illustrates the health of a firm in the short run. Whether it can settle its immediate debts using the available resources or not, this indicates that a decrease in WCTA means aggressive WCM. As described in the equation below (Berg et al., 2024).

$$\mathrm{WCTA}_{i,t} = \frac{\text{Current assets}_{i,t} - \text{Current liabilities}_{i,t}}{\text{Total assets}_{i,t}}$$

Cash Conversion Cycle (CCC): The CCC is one of the most commonly applied measures of WCM efficiency, which indicates the number of days it takes for a business to turn its investments in inventory and receivables into cash, minus the amount of time it gets credit from suppliers. The shorter the better, as it reflects more aggressive WCM, determined using Equation (b) (Liu et al., 2024).

$$\mathrm{CCC} = \mathrm{INV} + \mathrm{ACR} - \mathrm{ACP}$$

Inventory Turnover: $\mathrm{INV} = \frac{\text{Inventories}}{\text{Cost of goods}} \times 365$

Accounts receivable Turnover: $\mathrm{ACR} = \frac{\text{Accounts receivable}}{\text{Sales}} \times 365$

Accounts Payable Turnover: $\mathrm{ACP} = \frac{\text{Accounts payable}}{\text{Purchases}} \times 365$

Net Trade Cycle (NTC): The NTC captures the net time that a firm's operating resources spend stuck in the business and is an alternative measure of WCM. In contrast to the traditional Cash Conversion Cycle, NTC allows normalizing working capital elements by sales and expresses the outcome in terms of days, which makes it particularly useful when comparing firms across types and sizes. A shorter NTC implies more aggressive WCM, which is calculated using the following equation (Berg et al., 2024):

$$\mathrm{NTC} = \frac{\text{Inventories} + \text{Accounts receivable} - \text{Accounts payable}}{\text{Sales}} \times 365$$

3.2.1.3. Control variables

Following Liu and Jin (2024) and Sun and Gong (2023):

Logarithmic Change in Sales Revenue (ΔlnSALES): This variable represents the logarithmic change in sales between two periods, which can be interpreted as an indicator for the percentage change in sales. For cost stickiness models, this variable is often the main dependent variable explaining cost changes, typically (Anderson et al., 2003).

$$\Delta \ln SALSE = \ln SALES_{i,t} - \ln SALES_{i,t-1}$$

Decrease Dummy (D): is an indicator of decreased sales, which can be taken equal to one and zero in cases where the sales revenue of the current period is estimated as less when compared to that of the previous year and otherwise, respectively, revenue (Anderson et al, 2003).

Free cash flow (FCF): It is one of the important indicators in finance and accounting research that is considered in the framework of Agency Theory. According to this theory, managers may invest in low-value or low-return projects when they have free cash, and this can increase the conflict of interest between managers and owners. For this reason, many studies, including Chen et al. (2012), have used FCF as a key variable in analyzing managerial behavior (Chen et al., 2012).

$$FCF_{i,t} = \frac{Operating\ Cash\ FloW_{i,t} - Dividends_{i,t}}{Total\ Assets_{i,t}}$$

Employee intensity (EMINT) is measured by dividing the number of employees by sales. This indicator shows how many human resources the company has employed in relation to its sales volume and the intensity of labor utilization in the company's operational structure (Chen et al, 2012).

$$EMINT_{i,t} = \frac{Number\ Of\ Employees_{i,t}}{Total\ sales_{i,t}}$$

Firm size (SIZE): Firm size is calculated as the natural logarithm of total assets (Huynh et al., 2025).

Profitability (ROA): Profitability is measured as net income divided by total assets. ROA reflects operational performance and internal financing capacity (Huynh et al., 2025).

$$ROA_{i,t} = \frac{Net\ Income_{i,t}}{Total\ Assets_{i,t}}$$

Cash holding: Ratio of operating cash flow to total assets. This ratio measures the company's cash holdings based on cash flow from operations relative to total assets. This metric shows the company's operational liquidity (Habib and Hasan, 2019).

$$Cash\ Holding_{i,t} = \frac{operating\ cash\ floew_{i,t}}{Total\ Assets_{i,t}}$$

Changes in Gross Domestic Product (ΔGDP) are incorporated as a control variable to represent changes in Gross Domestic Product. This variable accounts for the effects of inflation and managerial expectations. (Banker and Byzalov, 2014).

$$\Delta \mathrm{GDP}_{\mathrm{i,t}} = \frac{\mathrm{GDP_t}}{\mathrm{GDP_{t-1}}} - 1$$

Model 2: Effect of WCM on financial constraints (Mediation)

The next model analyzes the mediating role of financial constraints with respect to the SA index (Hadlock and Pierce, 2010). The SA index is based on only exogenous variables (firm size and age) and is the least endogenously biased one compared to other indices (for example, KZ). The SA index was chosen based on the cost stickiness model and literature review background (Costa et al., 2021; Wang, 2022).

On the basis of prior empirical results, and in particular Huynh et al. (2025), the subsequent model has been applied to investigate the influence of working capital management on financial constraints.

$$SA_{i,t} = \beta_0 + \beta_1 WCM_{i,t} + \beta_2 Leverage_{i,t} + \beta_3 SIZE_{i,t} + \beta_4 Cash\ holding_{i,t} + \beta_5 ROA_{i,t} + \beta_6 Salesgrowth_{i,t} + \beta_7 \Delta GDP_{i,t} + \varepsilon_{i,t}$$

The regression equations are estimated three times by replacing WCM with CCC, NTC, and WCTA.

3.2.2. Variable definitions and measurement (Model 2)
3.2.2.1. Dependent variable

Hadlock–Pierce Financial Constraints Index (SA): One of the most reliable and non-parametric indicators for measuring the degree of financial constraint of firms was developed by Hadlock and Pierce (2010). This indicator focuses on the fundamental characteristics of the firm, because the authors believed that financial variables such as cash flow or profitability may themselves be the result of financial constraint and introduce bias. This indicator uses only firm size and firm age, because these two variables are less susceptible to endogenous bias than changes in financial policies. A higher value of this index indicates greater financial constraints.

$$\mathrm{SA_{i,t}} = -0.737\mathrm{SIZE_{i,t}} + 0.043\mathrm{SIZE_{i,t}^{2}} - 0.04\mathrm{AGE_{i,t}}$$

3.2.2.2. Independent variable

WCM is measured using three separate indicators; therefore, the model is estimated three times, once with CCC, once with NTC, and once with WCTA as the working capital proxy (Berg et al, 2024).

3.2.2.3. Control variables

Leverage: It shows the extent to which a company relies on debt. The leverage ratio shows what portion of a company's assets is financed through debt. Several studies, including Banerjee et al. (2023), emphasize that increasing leverage can increase debt repayment risk and exacerbate the company's level of financial constraints. Leverage is measured as total debt divided by total assets (Banerjee et al, 2023).

$$Leverage_{i,t} = \frac{Total\ Debts_{i,t}}{Total\ Assets_{i,t}}$$

Firm Size (SIZE): measured by total assets, the natural logarithm. SIZE is a prevalent control variable in empirical studies of corporate finance and accounting because larger firms are universally less subject to financing frictions and better situated to obtain external capital (Huynh et al., 2025).

Cash holding: Operating cash flow to total assets ratio. This ratio calculates the cash the company holds based on cash from operations divided by total assets. This metric indicates the operational liquidity of the company. ((Habib and Hasan, 2019).

$$Cash\ Holding_{i,t} = \frac{operating\ cash\ floew_{i,t}}{Total\ Assets_{i,t}}$$

Profitability (ROA): Profitability is measured as net income divided by total assets. ROA indicates operating performance and internal financing ability (Huynh et al., 2025).

$$ROA_{i,t} = \frac{Net\ Income_{i,t}}{Total\ Assets_{i,t}}$$

Sales Growth: Sales growth is defined as the change in sales from year t−1 to year t scaled by year t−1 sales. Higher sales growth increases a firm’s need for external financing and may intensify financial constraints, so this variable is a widely used proxy of growth-related financing pressure in corporate finance research (Anderson et al, 2003).

$$Sales\ Growth_{i,t} = \frac{SALES_{i,t} - SALES_{i,t-1}}{SALES_{i,t}}$$

Changes in Gross Domestic Product (ΔGDP) are incorporated as a control variable to represent changes in Gross Domestic Product. This variable accounts for the effects of inflation and managerial expectations. (Banker and Byzalov, 2014).

$$\Delta GDP_{i,t} = \frac{GDP_t}{GDP_{t-1}} - 1$$

3.2.3. Resulting structure

This specification clearly establishes that the dependent variable is financial constraints, the main independent variable is working capital management (measured through three alternative proxies), and all other variables function as standard control variables consistent with international empirical literature.

After examining the first model, which is the direct path of the relationship between WCM and cost stickiness according to the model of Chen et al. 2003 and Berg et al. 2024, the second model is examined by examining the indirect path that shows the relationship and type of effect of WCM on financial constraints, and confirmation of the two direct and indirect relationships is obtained. According to Baron and Kenny (1986), a mediating effect exists when the independent variable affects the mediator, and the mediator in turn affects the dependent variable. Then, to finally examine the SA of financial constraints in the model (1), we examine it with the structural equation mediation method and whether the indirect effect in the example is statistically significant, we can use bootstrapping or the Sobel Test (Sobel 1998). Bootstrapping is statistically more accurate and flexible, because it does not depend on the normality of the data, so the bootstrapping test has been used for high accuracy of the study.

4. Findings

4.1. Descriptive statistics

Table 1 shows the descriptive analysis of the variables used in this research.

Table 1. Descriptive statistics

Variable	Mean	Minimum	Maximum	Std.dev	Skewness	Kurtosis	N
ΔlnCOST	0.226	-3.160	4.090	0.365	0.376	23.810	1848
ΔlnSALES	0.227	-0.435	0.899	0.279	-0.016	2.680	1848
WCTA	0.171	-1.950	2.006	0.216	-0.220	3.470	1848
CCC	200.100	-2025	3101	190.900	-0.304	13.530	1848
NTC	233.500	-2376	2689	295.900	0.927	15.120	1848
SA	-1.660	-3.165	3.997	0.908	1.710	7.860	1848
FCF	0.024	-0.814	0.538	0.127	-0.775	7.025	1848
EMINT	0.005	1.900	0.008	0.000	3.740	26.540	1848
SIZE	14.480	10.030	21.570	1.611	0.665	4.050	1848
ROA	0.141	-0.400	0.681	0.144	0.677	3.890	1848
Cashholding	0.125	-0.460	0.726	0.140	0.539	4.140	1848
Leverage	0.355	0.003	0.953	0.220	0.380	2.140	1848
Salesgrowth	0.331	-0.739	3.770	0.452	1.87	10.700	1848
ΔGDP	- 0.480	-3.727	2.909	1.743	0.115	2.49	1848
D	0.192	0.000	1.000	0.000	1.550	3.420	1848

Table 1 presents the descriptive statistics of the variables used in this study. WCM is quantified using three proxies: WCTA, CCC, and NTC. The mean WCTA is 0.171 or approximately 17%, indicating that firms have good working capital and a good liquidity position. According to Eldomaty et al. (2023), a WCTA between 5% and 30% suggests appropriate working capital levels along with the adoption of aggressive WCM policies. The CCC median is 200 days, which means that firms take around 200 days to achieve the whole cycle from raw material cash outflows through to customer cash inflows. Likewise, the mean NTC is 233 days, which represents the sum of accounts payable period, inventory holding period, and days' sales in accounts receivable. The mean value of the decrease in sales variable is 0.1926, indicating that the firm experienced a year-over-year sales reduction in about 19% of the observations.

4.2. Main results

Table 2 presents the results of the first model's estimation under hypothesis testing H1. The adjusted R2, Fisher statistics, and p-values ($p < 0.001$) all confirm the joint significance of the regression models. The Hausman test was significant in every one of the three estimates ($p < 0.001$), in support of employing the fixed effects model. Additionally, both the LR test for heteroscedasticity and the Wooldridge test for autocorrelation were significant, so fixed effects with robust standard errors and clustered standard errors were used in an attempt to solve these issues. The mean VIF scores were below 3 on all three estimates, and below 4 in all variables, indicating no indication of multicollinearity. All three WCM proxies, CCC, NTC, and WCTA, have their estimated coefficients positive and significant at 5% level. The sign is positive, suggesting that the higher values of these WCM measures are associated with a reduction in cost stickiness. Thus, hypothesis H1 holds. This result supports previous research, including Filbeck and Krueger (2005) and Berg et al. (2024), because they found effective working capital management also lowers cost inflexibility, such that cash can better adapt. This makes sense as a positive and significant value of coefficients for WCTA, NTC and CCC suggests that an increase in the working capital management indicators leads to decreasing cost stickiness. The positivity of the interaction coefficient ($D \times \Delta \ln Sales \times WCM$) indicates that during periods of declining sales, costs decrease faster and, as a result, the cost structure becomes more flexible. This finding fully confirms the first hypothesis (H1). Theoretically, this result is also consistent with the Adjustment Cost Theory, because companies with lower levels of unused resources incur lower adjustment costs and have a greater ability to quickly reduce committed resources during times of decline. The magnitude of the coefficients (with a t-statistic of about 2 to 3 and a significance level of less than 5 percent) indicates that the effect is also significant from an economic perspective. From a managerial perspective, the results suggest that adopting an aggressive approach to working capital management enables the company to adjust costs more quickly and effectively when faced with declining sales.

To examine the second hypothesis, we first examine the second model (Table 3) to see if there is a significant association between the WCM approach and financial constraints. If found, that association is then examined for its mediating effect. The Fisher statistics and their significance levels ($p < 0.001$) indicate the joint significance of the regression models in each of the three specifications. Both times, the Hausman test results, which are significant at the 0.1% level, prefer the fixed effects specification. Due to the significant LR test for heteroscedasticity and the Wooldridge test for autocorrelation, we employ the GLS model to adjust for these. Average VIF values still remain below 3 for all three models and below 4 for individual variables, confirming the absence of multicollinearity. The $\beta 1$ coefficients of all three WCM measures (CCC, NTC, and WCTA) are positive and significant at the 1% level, indicating that more aggressive WCM is associated with higher financial constraints. As such, these results confirm the mediation path specified in Hypothesis 2. Moreover, the introduction of the interaction term between WCM and sales decline for cost stickiness management confirmed the sign and significance of all coefficients obtained for Hypothesis 2 (see Table 4). The moderating role of financial constraints is in accordance with the Adjustment Cost Theory (Anderson et al., 2003), which argues that liquidity constraints increase cost stickiness due to a lack of flexibility. The positive and statistically significant coefficients of aggressive policies in all three WCM indices (namely, WCTA, CCC and NTC) suggest that more aggressive working capital management policies tend to raise companies' financial constraints. A positive coefficient indicates that the reduction in liquidity and higher dependence on external financing increase the tightening of financial constraints. This outcome supports path A of the second hypothesis (H2).

Table 2. Testing the First Hypothesis (fixed effect vce (robust))

ΔlnCost		**WCM Proxy**								
		CCC			**NTC**			**WCTA**		
Sym	**Variables**	**Coef**	**T-value**	**Prob**	**Coef**	**T-value**	**Prob**	**Coef**	**T-value**	**Prob**
β_1	ΔlnSales	0.892	19.570	0.000	0.897	19.460	0.000	0.866	18.610	0.000
β_2	D*ΔlnSales	-0.020	-0.140	0.887	-0.019	-0.130	0.893	0.089	0.880	0.380
β_3	D*ΔlnSales* WCM	0.320	1.740	0.084	0.332	2.030	0.044	0.535	2.040	0.043
β_4	WCM	0.202	2.240	0.027	0.218	3.130	0.002	0.227	3.070	0.003
β_5	FCF	-0.682	-2.960	0.004	-0.681	-2.990	0.003	-0.592	-2.580	0.011
β_6	EMINT	-0.520	-2.410	0.018	-0.470	-2.280	0.024	-0.414	-2.130	0.035
β_7	SIZE	0.022	2.460	0.015	0.022	2.460	0.015	0.022	2.630	0.010
β_8	ROA	-0.110	-5.900	0.000	-0.109	-6.000	0.000	-0.12	-6.170	0.000
β_9	Cashholding	0.799	3.510	0.001	0.801	3.630	0.000	0.681	3.170	0.002
β_{10}	ΔGDP	0.0153	3.160	0.002	0.015	3.280	0.001	0.015	3.460	0.001
	Constant	YES			YES			YES		
	Firm control	YES			YES			YES		
	Industry control	YES			YES			YES		
	Year control	YES			YES			YES		
	Adjusted R^2	0.582			0.583			0.583		
	F-statistics	134.360			163.710			177.460		
	Prob	0.000			0.000			0.000		
	Hausman	46.930			51.910			50.870		
	Prob	0.000			0.000			0.000		
	Lr test	3180.520			3176.150			3169.110		
	Prob	0.000			0.000			0.000		
	Wooldridge	4.500			4.834			5.098		
	Prob	0.035			0.029			0.025		
	Mean VIF	2.250			2.24			2.270		

Table 3. Testing the SA and WCM relationship (GLS)

SA		WCM Proxy								
		CCC			NTC			WCTA		
Sym	Variables	Coef	Z-value	Prob	Coef	Z-value	Prob	Coef	Z-value	Prob
β_1	WCM	0.0382	4.330	0.000	0.035	3.990	0.000	0.0331	4.250	0.000
β_2	Leverage	0.023	2.720	0.007	0.0236	2.670	0.008	0.333	3.690	0.000
β_3	SIZE	0.503	341.400	0.000	0.503	3410	0.000	0.504	3430	0.000
β_4	Cashholding	-0.002	-0.100	0.918	-0.000	-0.040	0.970	-0.0035	-0.290	0.769
β_5	ROA	0.0215	1.430	0.154	0.023	1.590	0.112	0.014	0.900	0.366
β_6	Salesgrowth	0.0113	2.970	0.003	0.011	2.870	0.004	0.0073	1.980	0.048
β_7	ΔGDP	0.0097	12.790	0.000	0.009	12.840	0.000	0.0096	12.700	0.000
	Constant		YES			YES			YES	
	Firm control		YES			YES			YES	
	Industry control		YES			YES			YES	
	Year control		YES			YES			YES	
	Wald		154391			154058			154566	
	Prob		0.000			0.000			0.000	
	Hausman		78.180			76.740			80.680	
	Prob		0.000			0.000			0.000	
	Lr test		2961			2960			2958.460	
	Prob		0.000			0.000			0.000	
	Wooldridge		278.600			278.616			304.600	
	Prob		0.000			0.000			0.000	
	Mean VIF		1.480			1.470			1.590	

Theoretically, these items are completely in line with the liquidity-yield trade-off theory and Pecking Order Theory since a reduction in working capital may enhance liquidity risk and force the company to pay much more, depending on external financing. However, the CCC index coefficient was found to be the largest among all five treatment variables of time-mismatch liquidities, suggesting that the liquidity time cycle has a greater impact on financial constraints. From a managerial perspective, companies that reduce working capital too much should expect to become more dependent on costly and difficult financing during periods of economic downturn or volatility.

The covariance-based structural equation model (SEM), estimated by maximum likelihood, provides the result of the second hypothesis as shown in Table 4. The results of the confirmatory factor analysis (CFA), presented in the bottom panel of the table, indicate a good fit of the model, as it meets the following goodness-of-fit criteria (RMSEA < 0.08, CFI > 0.90, TLI > 0.90, and SRMR < 0.05).

The SEM results show that aggressive WCM has both a direct and an indirect effect on cost stickiness. The significant negative path WCM → SA indicates that these policies increase financial constraints. Also, the significant negative path SA → cost stickiness indicates that higher financial constraints strengthen cost stickiness. Therefore, although aggressive WCM directly reduces cost stickiness (efficiency channel), it can increase stickiness through increasing financial constraints (vulnerability channel). This finding supports the hypothesis that there is partial mediation, as outlined in Hypothesis 2. This regularity fits the theory of adjustment costs, in that liquidity constraints preclude firms from realizing fast cost adjustments. From a manager's perspective, the finding suggests that aggressive activities toward WCM policies are desirable, provided the firm has reasonable access to financial resources or credit lines; otherwise, such tendencies for alleviating cost stickiness may be counterbalanced by financial pressures.

Table 5 presents the outcome of the bootstrapping mediation analysis (MacKinnon et al., 2004), a nonparametric resampling method that eliminates the normality assumption inherent in the Sobel test by empirically estimating the sampling distribution. The indirect effects were calculated with 2,000 bootstrap resamples. Bootstrap results show that the indirect effects of all three WCM proxies (CCC, NTC, and WCTA) on cost stickiness through the SA index are significant at the 95% confidence level. Since neither the minimum nor the maximum confidence interval includes zero, mediation is definitely confirmed. The mediation for the CCC index is stronger than for the other indices, indicating that the timing of the cash cycle is the most important channel of the effect of WCM on cost stickiness. These results fully support the second hypothesis and highlight the key role of financial constraint in cost adjustment behavior.

Table 4. SEM Results- WCM and cost stickiness mediated by SA index

Mediated SEM	WCM Proxy								
		CCC			NTC			WCTA	
Path	**Coef**	**Z-value**	**Prob**	**Coef**	**Z-value**	**Prob**	**Coef**	**Z-value**	**Prob**
WCM→SA Index	-1.468	-25.890	0.000	-1.074	-18.840	0.000	-1.21	-9.650	0.000
SA ndex→Cost sticky	-0.639	-17.370	0.000	-0.658	-19.110	0.000	-0.671	-20.810	0.000
WCM→Cost sticky	0.334	3.190	0.001	0.277	3.010	0.003	0.717	4.020	0.000
Controled		YES			YES			YES	
Chi-square χ^2		26.950			29.150			37.150	
RMSEA		0.037			0.040			0.048	
CFI		0.985			0.980			0.968	
TLI		0.962			0.951			0.921	
SRMR		0.025			0.026			0.031	
N		1848			1848			1848	

Table 5. Bootstrap mediation test

Mediation	Estimate	Lower	Upper	Prob	Result Mediation
CCC→SA→ Cost sticky	0.939	0.811	0.999	0.000	Supported
NTC→SA→Cost sticky	0.707	0.604	0.810	0.000	Supported
WCTA→SA→Cost sticky	0.813	0.631	0.995	0.000	Supported

Overall, the findings lend support to both hypotheses, which is that the relationship between WCM strategies and cost behavior is significant in both direct and indirect (through financial constraints) ways. The results strengthen the theoretical premise that liquidity decisions by management influence cost responsiveness and provide partial support for the Adjustment Cost Theory in an emerging market environment.

5. Conclusion

The findings delivered robust support for the study's core hypothesis; sets of working capital management that are aggressive exhibit lower cost stickiness, but their beneficial effect is partial because they also lead to higher levels of financial frictions. Not only does this dual effect confirm hypothesis H1, but it also supports the partial mediation shown by hypothesis H2. The positive and significant interaction term (D × ΔlnSales × WCM) also suggests that firms with aggressive WCM are reacting faster to the decline in sales, confirming the premise that WCM serves as a driver of operational flexibility. To be more specific, when a company suffers a 10% decrease in sales, firms with high (versus low) WCM are able to reduce their operating costs by 2–3% more than they would otherwise (and doing so within the confines of its working capital policy), demonstrating that aggressive levels of working capital are not just quantitative concepts but has direct application in simple operational settings?

Incorporating financial constraints into the model adds to recent calls in the literature (Kayani et al., 2019; Jabbouri et al., 2024) for integrating contextual factors that drive cost behavior. The results also indicated that while aggressive WCM improves cost adaptability, the extent of this improvement is contingent upon the firm's financial flexibility. The positive relationship between aggressive WCM and financial constraints is in line with the pecking-order theory and reflects liquidity risk existing in aggressive WCM policies. Partial mediation through SEM and bootstrap analysis suggests that financial constraints are not only a result of liquidity decisions but represent an active course of action driving the direction of cost adjustment. One caveat is that these results are observed in an emerging market with relatively high inflation, credit constraints, and an underdeveloped capital market. So, the extent to which working capital management will reduce cost stickiness could differ owing to institutional factors like restricted access to external financing and macroeconomic volatility, and thus, we need to consider contextual characteristics while explaining these results.

The findings expand WCM research, which traditionally focused on profitability and liquidity, to newly developed structural retention costs for asymmetric resource adjustments. By linking these two streams, the results clarify the role of liquidity management vis-à-vis adjustment-cost mechanisms in determining cost rigidity. That builds on earlier work like Lyngstadaas (2020), Baños-Caballero et al. (2013), Anderson and Lanen (2007) and Chen et al. (2012).

From a managerial perspective, the findings warned against implementing aggressive WCM strategies without factoring in their financial by-products. While managers might gain/cost flexibility in downturns due to better cost-pass-throughs, they can only do so if the firm has sufficient access to external finance/long-term liquidity buffers that nullify the clutches of financial constraints. Firms without these buffers may be subject to jitters that create additional financial friction, dissipating the net benefits in terms of cost responsiveness. This implies that aggressive WCM strategies yield positive effects only if enough financial resources or liquidity reserves are present.

So balancing efficiency gains with resilience becomes necessary.

The findings underscore the need for targeted liquidity-support mechanisms (particularly in the context of SMEs), whereas policymakers should heed them by emphasizing targeted liquidity-support mechanisms. And, since enhanced credit access can also mitigate cost stickiness (Banerjee and Duflo, 2005), firms are likely to reap the full benefit of aggressive WCM strategies in times of economic turmoil.

In sum, this study reinterpreted WCM not only as a liquidity instrument but also as an emerging strategic enabler of cost flexibility, whose effectiveness is contingent on the financial health of the firm. These findings supported the view that while WCM can improve operational flexibility, this is context-specific, especially in emerging markets characterized by strong institutional and macroeconomic constraints. Subsequent research might expand these insights by studying such cross-industry differences or examining the extent to which factors that determine manager behavior may have implications for the efficiency versus adaptability trade-off when liquidity is constrained.

References

1. Agegnew, A. (2019). The effect of working capital management on profitability: The case of selected manufacturing and merchandising companies in Hawassa city administration. *Research Journal of Finance and Accounting*, 10(1), pp. 51-85. https://doi.org/10.7176/RJFA/10-1-07
2. Aktas, N., Croci, E. and Petmezas, D. (2015). Is working capital management value-enhancing? Evidence from firm performance and investments. *Journal of Corporate Finance*, 30, pp. 98-113. https://doi.org/10.1016/j.jcorpfin.2014.12.008
3. Anderson, M. C., Banker, R. D. and Janakiraman, S. N. (2003). Are selling, general, and administrative costs "sticky"? *Journal of Accounting Research,* 41(1), pp. 47-63. https://doi.org/10.1111/1475-679X.00095

4. Anderson, M., Banker, R., Huang, R. and Janakiraman, S. (2007). Cost behavior and fundamental analysis of SG&A costs. *Journal of Accounting, Auditing & Finance*, 22(1), pp. 1-28. https://doi.org/10.1177/0148558X0702200103
5. Anderson, S. W. and Lanen, W. N. (2007). Understanding cost management: what can we learn from the evidence on'sticky costs'? *University of California, California, Berkeley*. Available at SSRN 975135. https://dx.doi.org/10.2139/ssrn.975135
6. Atanasova, C. and Willeboordse, F. H. (2025). The impact of corporate diversification on liquidity management: Evidence from lines of credit. *Journal of International Financial Markets, Institutions and Money*, 101, A.102139. https://doi.org/10.1016/j.intfin.2025.102139
7. Ballas, A., Naoum, V. C. and Vlismas, O. (2022). The effect of strategy on the asymmetric cost behavior of SG&A expenses. *European Accounting Review,* 31(2), pp. 409-447. https://doi.org/10.1080/09638180.2020.1813601
8. Banerjee, S., Mitra, A. and Mohanti, D. (2023). Leverage and corporate investment–a cross-country analysis. *Investment Management & Financial Innovations*, 20(3), pp. 126. https://doi.org/10.21511/imfi.20(3).2023.11
9. Banker, R. D. and Byzalov, D. (2014). Asymmetric cost behavior. *Journal of Management Accounting Research,* 26(2), pp. 43-79. https://doi.org/10.2308/jmar-50846
10. Banker, R. D. and Chen, L. (2006). Predicting earnings using a model based on cost variability and cost stickiness. *The Accounting Review*, 81(2), pp. 285-307. https://doi.org/10.2308/accr.2006.81.2.285
11. Baños-Caballero, S., García-Teruel, P. J. and Martínez-Solano, P. (2014). Working capital management, corporate performance, and financial constraints. *Journal of Business Research,* 67(3), pp. 332-338. https://doi.org/10.1016/j.jbusres.2013.01.016
12. Baron, R. M. and Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), A. 1173. https://doi.org/10.1037/0022-3514.51.6.1173
13. Berg, T., Gustafsson, E. and Wahlstrøm, R. R. (2024). Cost management and working capital management: ebony and ivory in perfect harmony? *Journal of Management Control*, 35(2), pp. 207-233. https://doi.org/10.1007/s00187-024-00368-3
14. Bottazzi, G., Secchi, A. and Tamagni, F. (2014). Financial constraints and firm dynamics. *Small Business Economics*, 42(1), pp. 99-116. https://doi.org/10.1007/s11187-012-9465-5
15. Chen, C. X., Lu, H. and Sougiannis, T. (2012). The agency problem, corporate governance, and the asymmetrical behavior of selling, general, and administrative costs. *Contemporary Accounting Research*, 29(1), pp. 252-282. https://doi.org/10.1111/j.1911-3846.2011.01094.x
16. Chen, Y. and Ma, Y. (2021). Financing constraints, internal control quality and cost stickiness. *Business Economics and Management (JBEM),* 22(5), pp. 1231-1251.
17. Chow, C. K. W. and Fung, M. K. Y. (2000). Small businesses and liquidity constraints in financing business investment: evidence from Shanghai's manufacturing sector. *Journal of Business Venturing,* 15(4), pp. 363-383. https://doi.org/10.1016/S0883-9026(98)00014-7
18. Costa, M. D., Habib, A. and Bhuiyan, M. B. U. (2021). Financial constraints and asymmetric cost behavior. *Journal of Management Control*, 32(1), pp. 33-83. https://doi.org/10.1007/s00187-021-00314-7
19. Deari, F., Kukeli, A., Barbuta-Misu, N. and Virlanuta, F. O. (2024). Does working capital management affect firm profitability? Evidence from European Union countries. *Journal of Economic and Administrative Sciences*, 40(5), pp. 1060-1080. https://doi.org/10.1108/JEAS-11-2021-0222

20. Deloof, M. (2003). Does working capital management affect profitability of Belgian firms? *Journal of Business Finance & Accounting*, 30(3-4), pp. 573-588. https://doi.org/10.1111/1468-5957.00008
21. Deneckere, R. J. and Kovenock, D. (1996). Bertrand-Edgeworth duopoly with unit cost asymmetry. *Economic Theory*, 8(1), pp. 1-25. https://doi.org/10.1007/BF01212009
22. Filbeck, G. and Krueger, T. M. (2005). An analysis of working capital management results across industries. *Mid-American Journal of Business*, 20(2), pp. 11-20. https://doi.org/10.1108/19355181200500007
23. Frankel, R., Levy, H. and Shalev, R. (2017). Factors associated with the year-end decline in working capital. *Management Science*, 63(2), pp. 438-458. https://doi.org/10.1287/mnsc.2015.2351
24. Gitman, L. J. (1974). Estimating corporate liquidity requirements: a simplified approach. *Financial Review,* 9(1), pp. 79-88. https://doi.org/10.1111/j.1540-6288.1974.tb01453.x
25. Habib, A. and Hasan, M. M. (2019). Corporate social responsibility and cost stickiness. *Business & Society*, 58(3), pp. 453–492. https://doi.org/10.1177/0007650316677936
26. Hadlock, C. J. and Pierce, J. R. (2010). New evidence on measuring financial constraints: Moving beyond the KZ index. *The Review of Financial Studies*, 23(5), pp. 1909-1940.https://www.jstor.org/stable/pdf/40604834.pdf
27. Hayes, A. F. and Scharkow, M. (2013). The relative trustworthiness of inferential tests of the indirect effect in statistical mediation analysis: Does method really matter? *Psychological Science,* 24(10), pp. 1918-1927. https://doi.org/10.1177/0956797613480187
28. He, J., Tian, X., Yang, H. and Zuo, L. (2020). Asymmetric cost behavior and dividend policy. *Journal of Accounting Research,* 58(1), pp. 989–1021. https://doi.org/10.1111/1475-679X.12328
29. Hung, N. T. and Su Dinh, T. (2022). Threshold effect of working capital management on firm profitability: evidence from Vietnam. *Cogent Business & Management,* 9(1), A. 2141090. https://doi.org/10.1080/23311975.2022.2141090
30. Huynh, T. X. T., Nguyen, T. T. H. and Nguyen, C. V. (2025). The impact of working capital management on the financial performance of listed enterprises: an empirical evidence from Vietnam. *Cogent Business & Management*, 12(1), A. 2473033. https://doi.org/10.1080/23311975.2025.2473033
31. Ibrahim, A. E. A., Ali, H. and Aboelkheir, H. (2022). Cost stickiness: A systematic literature review of 27 years of research and a future research agenda. *Journal of International Accounting, Auditing and Taxation*, 46, A. 100439. https://doi.org/10.1016/j.intaccaudtax.2021.100439
32. Jabbouri, I., Satt, H., El Azzouzi, O. and Naili, M. (2024). Working capital management and firm performance nexus in emerging markets: do financial constraints matter? *Journal of Economic and Administrative Sciences*, 40(5), pp. 1020-1030. https://doi.org/10.1108/JEAS-01-2022-0010
33. 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. https://doi.org/10.1002/9780470752135.ch17
34. Jose, M. L., Lancaster, C. and Stevens, J. L. (1996). Corporate returns and cash conversion cycles. *Journal of Economics and Finance*, 20(1), pp. 33–46. https://doi.org/10.1007/BF02920497
35. Kayani, U. N., De Silva, T. A. and Gan, C. (2019). A systematic literature review on working capital management–an identification of new avenues. *Qualitative Research in Financial Markets,* 11(3), pp. 352-366. https://doi.org/10.1108/QRFM-05-2018-0062

36. Liu, L., Zhou, X. and Xu, J. (2024). Does working capital management improve financial performance in China's agri-food sector during COVID-19? A comparison with the 2008 financial crisis. *Plos One*, 19(4), A. e0300217. https://www.emerald.com/mf/article/50/3/534/1235493/The-impact-of-working-capital-management-on-the
37. Lyngstadaas, H. (2020). Packages or systems? Working capital management and financial performance among listed U.S. manufacturing firms. *Journal of Management Control*, 31, pp. 403–450. https://doi.org/10.1007/s00187-020-00306-z
38. Lyngstadaas, H. and Berg, T. (2016). Working capital management: Evidence from Norway. *International Journal of Managerial Finance*, 12, pp. 295–313. https://doi.org/10.1108/IJMF-01-2016-0012
39. Mirón Sanguino, Á. S., Crespo-Cebada, E., Muñoz, E. M. and Caro, C. D. (2024). Working capital: Development of the field through scientific mapping: An updated review. *Administrative Sciences*, 14(4), p. 67. https://doi.org/10.3390/admsci14040067
40. Prasad, P., Narayanasamy, S., Paul, S., Chattopadhyay, S. and Saravanan, P. (2019). Review of literature on working capital management and future research agenda. *Journal of Economic Surveys,* 33(3), pp. 827-861. https://doi.org/10.1111/joes.12299
41. Restuti, M. D., Gani, L., Shauki, E. R. and Leo, L. (2022). Does managerial ability lead to different cost stickiness behavior? Evidence from ASEAN countries. *International Journal of Financial Studies*, 10(3), p. 48. https://doi.org/10.3390/ijfs10030048
42. Singh, H. P., Kumar, S. and Colombage, S. (2017). Working capital management and firm profitability: A meta-analysis. *Qualitative Research in Financial Markets*, 9, pp. 34–47. https://doi.org/10.1108/QRFM-06-2016-0018
43. Ujah, N. U., Tarkom, A. and Okafor, C. E. (2020). Working capital management and managerial talent. *International Journal of Managerial Finance*, 17, pp. 455–477. https://doi.org/10.1108/IJMF-12-2019-0481
44. Vlismas, O. (2024). The moderating effects of strategy on the relation of working capital management with profitability. *Journal of Accounting & Organizational Change*, 20(2), pp. 276-306. https://doi.org/10.1108/JAOC-01-2023-0005
45. Wang, Z. (2022). Digital finance, financing constraint and enterprise financial risk. *Journal of Mathematics,* 2022(1), A. 2882113. https://doi.org/10.1155/2022/2882113

IJAAF; Vol. 10, No. 2, Spring 2026, Serial 35

Ferdowsi University of Mashhad

Iranian Journal of
Accounting, Auditing & Finance

Quarterly

RESEARCH ARTICLE

Audit Fees under Dual Earnings Management: The Moderating Role of Accruals in the REM-Audit Fee Nexus

Seyed Hamid Jalali, Hassan Hemmati
Department of Accounting, Kashan Branch, Islamic Azad University, Kashan, Iran
Omid Faraji*
Accounting and Financial Management Department, College of Farabi, University of Tehran, Qom, Iran
Hossein Jabbari
Department of Accounting, Kashan Branch, Islamic Azad University, Kashan, Iran

How to cite this article:
Jalali, S. H. , Hemmati, H. , Faraji, O. and Jabbari, H. (2026). Audit Fees under Dual Earnings Management: The Moderating Role of Accruals in the REM-Audit Fee Nexus. Iranian Journal of Accounting, Auditing and Finance, 10(2), 171-195. doi: 10.22067/ijaaf.2026.47628.1595
https://ijaaf.um.ac.ir/article_47628.html

ARTICLE INFO

Article History
Received: 2024-07-11
Accepted: 2025-11-03
Published online: 2026-04-12

Abstract

As financial reporting becomes increasingly complex, both real earnings management (REM) and accrual-based earnings management (AEM) pose challenges for auditors in risk assessment and fee determination. While prior studies have examined these strategies in isolation, scant attention has been paid to their concurrent use, dual earnings management, particularly in audit pricing. Existing literature lacks analyses of the REM–AEM interaction within moderating frameworks, leaving unclear how AEM intensity moderates the REM–audit fee relationship. This gap impedes understanding of audit pricing in information-asymmetric settings and may lead to suboptimal audit judgments. To address this gap, this study investigated how AEM moderates the relationship between REM and audit fees. The study offered three novelties. First, it is the first in the domestic literature to examine the REM–AEM interaction. Second, it treated AEM as a moderator, revealing how auditors respond to compounded informational risk. Third, it provided evidence from the Iranian capital market, offering a practical framework for audit risk assessment in low-transparency environments. Using data from 72 firms listed on the Tehran Stock Exchange (2013–2025), results indicated a positive and significant relationship between REM and audit fees. A negative and significant interaction between REM and AEM emerges only when AEM is measured by the Kothari model. Results showed that using both strategies simultaneously affects auditors' risk perceptions and weakens REM's direct impact on audit fees. The study contributed to audit pricing theory by bridging behavioral and signaling perspectives, and presents practical implications for auditors as well as regulators and standard-setters.

Keywords:
Accrual Earnings Management, Audit Fees, Dual Earnings Management, Real Earnings Management, Moderating Role

https://doi.org/10.22067/ijaaf.2026.47628.1595

CC BY

NUMBER OF REFERENCES
46

NUMBER OF FIGURES
0

NUMBER OF TABLES
10

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

*Corresponding Author: Omid Faraji
Email: omid_faraji@ut.ac.ir
Tel: 09126515510
ORCID: 0000-0001-7678-3126

1. Introduction

The link between earnings management and audit fees is a mainstay of accounting research for decades, with the relationship having received considerable attention within the literature (Ashbaugh et al., 2003; Frankel et al., 2002; Schroeder, 2002). As per ReferenceSchipper (1989), earnings management refers to a purposeful intervention in the financial reporting process, with the intent of obtaining some predetermined target earnings. Such intervention can take the form of income-increasing, -decreasing, or -smoothing behaviors. Earnings management would be a general term, but it can be categorized into two types, namely accrual-based earnings management (AEM) and real earnings management (REM), where AEM refers to the manipulation of accounting accruals and estimates, while REM is related to operational actions (Choi et al., 2022). When both are used concurrently, the practice is known as dual earnings management. Accumbent complexity in the process of auditing could skew the risk assessment, coupled with the authenticity of reporting by auditors.

AEM and REM both impact audit pricing directly, as they increase perceived audit risk and complexity. AEM, in particular, raises the audit fees due to higher accounting estimation risk and potential misstatement, whereas REM increases the audit fees owing to operational and business risk, which requires more extensive audit effort. Thus, the interplay of these two strategies can lead to substantial changes in auditors' risk assessments and pricing decisions. Evidence from previous studies also indicates that both accrual-based and REM have the potential to lower the quality of financial reporting and affect auditors' readers about subsequent risk (Sun et al., 2014). Such factors together create informational asymmetry in the context of earnings management that increases audit risk (Choi et al., 2024; Kim and Park, 2014; Krishnan et al., 2013; Kyriakou, 2025), which could increase audit fees. Thus, for practitioners and regulatory entities alike, it is essential to develop a more in-depth knowledge regarding the impact of different forms of earnings management on audit fees.

The link between REM and AEM, the two most common kinds of earnings management, and audit pricing has received less research than the relationship between earnings management and audit fees. The majority of previous research has examined these management techniques independently, without considering their interplay or relative impact on audit fee structures. For instance, different studies have looked at how REM or AEM affect audit fees separately (Almarayeh et al., 2020; Azad et al., 2023; Choi et al., 2022; Egbunike et al., 2023; Emadoddini and Saeedi, 2019; Hussien et al., 2024; Sitanggang et al., 2020), but there is not much information regarding how auditors react when both methods are used at once. Due to this lack of clarity, auditors may have a more difficult time doing risk assessments and estimating audit effort accurately.

Specifically, there has been a dearth of studies that employ moderating frameworks to conduct interactive analyses between REM and AEM, which is the root cause of the current knowledge gap. Put simply, the impact of AEM intensity on the correlation between REM and audit fees is still not well understood. As a result, auditors may make less-than-ideal decisions due to a lack of understanding of audit pricing mechanisms in information-asymmetric settings.

The study seeks to fill this gap by examining the moderating role of discretionary accruals as a proxy for AEM on post-audit fees. In particular, it investigates whether discretionary accruals weaken or strengthen the relationship between REM and audit pricing. Utilizing a quantitative approach and an empirical modeling based methodology, the study investigates the collective effect of these two managerial behaviours on audit fee setting. The analysis focuses on the case where firms dual engage in REM and AEM, which forces their auditors to update (or better) adjust for potentially contradictory information provided by managers.

The novelty of this study lies in three key dimensions. This study is, to the best of the author's knowledge, the first in domestic literature that systematically ponders the interactive role of REM and AEM in contrast with previous literature, where literature has mostly treated them separately. Second, it provides a fresh view for auditors' reactions to the increased informational risk by suggesting AEM as a moderating variable. Third, by presenting empirical evidence from the Iranian capital market, the study provides a practical framework for assessing audit risk and estimating audit fees in environments characterized by low financial transparency. More broadly, all the contributions together will help develop a theoretical understanding of interactions between earnings management and audit fees.

Finally, answering this research question can provide useful insights for regulation-oriented organizations such as the Securities Exchange Organization and Audit Organization in Iran, as well as the Iranian Association of Certified Public Accountants. Insights into AEM's potential moderating role in the REM and audit fees relationship can aid in creating more focused audit pricing models, as well as aiding the identification of high-risk managerial behavior of firms that need increased scrutiny/intervention, helping enhance audit quality in the capital market.

2. Literature review and hypothesis development

2.1. Audit fees: theoretical underpinnings and core determinants

Audit fees are an important outcome variable in the audit pricing literature as they denote the economic remuneration auditors receive for bearing professional risk, allocating resources, and providing assurance over financial statements (Simunic, 1980). The seminal work of Simunic (1980) demonstrated that audit fees are primarily a function of firm size, business complexity, and inherent risk; factors that collectively drive the scope and intensity of audit effort. Subsequent research has expanded this framework by incorporating agency conflicts, information asymmetry, and managerial opportunism as key determinants of audit pricing (Abbott et al., 2006; Sohn, 2011). From the perspective of agency theory (Meckling and Jensen, 1976), external audits serve as a governance mechanism to align managerial behavior with shareholder interests. When managers engage in opportunistic actions, such as earnings manipulation, to mask poor performance or secure private benefits, auditors respond by intensifying procedures, increasing testing, and demanding higher compensation for elevated litigation and reputational risk (Greiner et al., 2017). Simultaneously, the theory of information asymmetry (Scott 2015) suggests that managers suffering from superior private information leads to uncertainty related to decisions made by outside stakeholders. In these circumstances, auditors serve as intermediaries between providers and users of information; the fees they charge increase with the incompleteness or opacity of financial reporting. Thus, this theory helps to guide our studies on how particular earning management forms (AEM & REM) affect audit fees. Both approaches represent managerial efforts to manage reported earnings; however, they differ significantly in terms of implementation, salience, and perceived risk, which reflects back into divergent, and possibly interrelated, effects on audit pricing.

2.2. AEM and its impact on audit fees

AEM essentially refers to the manipulation of accounting accruals, estimates or discretionary judgment available within the flexibility allowed by the relevant accounting standards (Jones, 1991). Although such practices may be technically in accordance with regulations, they can mask the economic reality and signal potential underlying financial distress or opportunism (Healy and Wahlen, 1999). According to signaling theory (Spence, 1978), abnormal accruals act as negative signals that make it harder to see through the finances and make auditors more skeptical. Empirically, AEM has been consistently linked to higher audit fees. Abbott et al. (2006) found that the direction of accrual manipulation is critical, such that income-increasing AEM in high-litigation environments elicits meaningful fee premia. In a study based in the UK, Alhadab (2018) suggested that although good quality auditors place strong constraints on AEM, it still can be constraining as the existence of abnormal accruals without any evidence about the direction raises audit fees reflecting auditors' risk-averse behaviour. Working with data up to October 2023, Yu (2020) demonstrated that non-specialized auditors have a unique sensitivity toward AEM; they charge bigger fees when the industry is not their speciality, decreasing reliance on skills in this area and increasing uncertainty, which justifies paying increased fees for eliminating those risks. More recently, Choi et al. whereas "big bath" accounting is a form of aggressive AEM during restructuring (2024), which has been associated with substantial risk to the audit process and potential future restatement, leading to significantly higher audit fees due to that informational as well as re-statement risk. Similarly, Hussien et al. (2024) reported that in Jordan, the engagement of high-quality auditors mitigates AEM, but only when institutional oversight is strong. This implies that in weak governance settings, AEM remains a potent driver of audit pricing. In the case of Iran, Darogheh Hazrati and Pahlevan (2013) were among the first to document a positive association between earnings management (mainly AEM) and audit fees. Emadoddini and Saeedi (2019) confirmed the relationship and revealed that the rise in abnormal audit fees is positively correlated with both AEM and REM. Khoshkar et al. (2019) noted that poorer earnings quality, commonly proxied by AEM, corresponds with fee increases as the auditor's forward-looking risk assessment. Together, these studies are evidence that AEM is not just a technical accounting choice; it is also a strategic signal that auditors perceive as greater reporting risk and something for which they should be compensated at higher rates.

2.3. REM and its distinct risk profile

REM is based on real operational decisions, unlike AEM, and includes reducing research and development or ad spend, overproducing inventory to reduce costs per unit or selling with abnormal discount (Roychowdhury, 2006). Since these actions influence cash flows and economic substance, REM is often considered to be more expensive and potentially harmful for long-term firm value (Cohen et al., 2010). Moreover, REM is ineluctably more difficult to detect as it overlaps with legitimate business choices and thus heightens audit risk (Choi et al., 2018). Early empirical evidence that REM increases audit costs, even after controlling for AEM, is provided by Sohn (2011), who concludes that auditors regard it as an independent risk. Greiner et al. 2017 framed "business risk" as the driver of REM and not bucketing it purely as an accounting issue, arguing that auditors would adjust their fees to reflect concerns over future solvability or operational sustainability. Choi et al. (2018) reinforced this view, showing that REM exerts an incremental effect on audit fees beyond AEM, particularly in firms with weak governance. A more nuanced view was provided by Sitanggang et al. (2020), who recorded mixed effects of REM components. They found that audit fees went up when discretionary spending went abnormally low, but audit fees went down when operating cash flows went abnormally high, which could be because of the impression of liquidity strength. This illustrates why REM measures need to be disaggregated when evaluating audit risk. In emerging markets, the REM–audit fee link is even more pronounced. Egbunike et al. (2023) studied Nigerian firms and found a negative relationship between residual audit fees and real earnings smoothing, interpreting this as evidence of intensified auditor oversight that suppresses opportunistic behavior. The indirect positive link between EM and audit fees was verified by Kyriakou (2025) in a cross-country analysis during economic crises. This association was mediated by heightened audit risk, particularly for REM, which tends to soar during downturns. There is strong evidence within Iran to back this up. Abbaszadeh et al. (2017) used R&D cuts as a proxy for REM and found a significant positive link with audit fees, consistent with the notion that operational manipulation raises red flags.

Moradi et al. (2019) provided evidence that REM leads to a rise in audit fees of TSE firms and claim that this is due to low transparency and weak monitoring mechanisms. Research in this area has shown that auditors can identify and charge for this kind of manipulation; for example, Emadoddini and Saeedi (2019) found that REM and AEM independently predict abnormal audit fees, and Azad et al. (2023) found a negative association between abnormal fees and smoother operating cash flows and production costs, which may indicate REM. Taken together, these results highlight REM as a key factor influencing audit pricing, especially in contexts where information is scarce, such as Iran.

2.4. The concurrent use of REM and AEM: Dual earnings management and audit risk complexity

An emerging body of literature acknowledges that firms do not typically exclusively employ one strategy of earnings management. Instead, managers often resort to a practice termed dual earnings management (the joint utilization of REM and AEM) that would allow them to satisfy earnings targets while at the same time reducing detection risk (Cohen and Zarowin, 2010). This practice further exacerbates informational asymmetry and creates complex risk profiles that traditional audit models cannot capture. Conceptually, the interaction of REM and AEM components may be one of substitution or distraction rather than being additive. Signaling theory suggests that salient AEM may dominate auditors' attention, causing them to underweight subtler operational red flags (REM), suggesting that when confronted with multiple risk indicators, practitioners tend to prioritize those that are more salient or familiar, typically accrual-based measures, while paying less attention to real activities, which remains underexplored. While Choi et al. (2018) did note that REM and AEM had additive effects, they failed to account for potential moderating effects. Although there may be interdependencies between the two methodologies, Sohn (2011) and Greiner et al. (2017) handled them separately. Not even newer studies, such as Choi et al. (2022), which found auditors respond more to REM than AEM, did not consider the effect of one on pricing the other. The gap is sharper still in Iran. Although research by groups like Emadoddini and Saeedi (2019) and Moradi et al. (2019) has shown that REM and AEM each have their own effects, no studies have attempted to model the interaction between the two. Most studies have been conducted within the country use linear regression models that disregard possible non-linearities or moderation in favor of additive effects (Darogheh Hazrati and Pahlevan, 2013; Khoshkar et al., 2019). Even Azad et al. (2023), whose focus was not on how AEM would act to condition auditors' responses, nor indeed did they define the REM proxies they were themselves considering. For instance, Alavi Tabari and Haji Moradkhani (2015) and Haiderpoor and Jafari (2016) examined the audit pricing factors determining in Iran; however, they paid considerable attention to market structure or liquidity but ignored features that vary depending on earnings management. Analyzing risk committees, Kordsholi and Pajoohi (2024) found no significant effect on audit fees or reporting quality and demonstrated that attention to managerial behavior is needed. Is the fee response dampened by auditors' distraction caused by high AEM, or does it become even more pronounced? Given the potential for auditor independence restrictions, restricted transparency, and concentrated ownership in Iran to worsen cognitive biases and risk mispricing, this is an important question to ask.

The following section synthesizes the reviewed literature and outlines the hypotheses that form the basis of our empirical investigation.

2.5. Hypothesis development

Based on the theoretical and empirical foundations discussed earlier, we are now ready to make our guesses about how REM, AEM, and audit fees are related. We begin by suggesting that audit fees are positively correlated with REM, and then we postulate that AEM acts as an attention-diverting signal to mitigate this correlation.

2.5.1. Direct effect: REM and Audit Fees

The likelihood of future financial trouble is increased because REM affects operational efficiency and puts substantial economic costs on enterprises (Roychowdhury, 2006; Cohen et al., 2010). Beyond simple financial manipulation, REM poses a risk to businesses. Greiner et al. (2017) noted that auditors see REM as a sign of aggressive management behavior that could threaten the company's long-term sustainability, which makes them increase charge premiums and conduct more rigorous testing. If there is little external oversight, market players may have a harder time spotting operational manipulations (Moradi et al., 2019). A large body of empirical research backs up this position. Auditors should consider REM as a separate risk factor, as Sohn (2011) discovered that it raises audit fees regardless of whether AEM is controlled for or not. In a similar vein, Choi et al. (2018) showed that REM affects audit pricing incrementally, especially in poorly governed environments. Findings from the Iranian studies by Emadoddini and Saeedi (2019) and Abbaszadeh et al. (2017) corroborate the positive correlation between audit fees and REM proxies, such as unusual reductions in research and development spending or increases in output, suggesting that auditors pay closer attention to unusual business practices. As a result, auditors should factor REM into their fee structures, in line with agency theory and the risk-based audit model. Therefore, we propose:

H1: REM is positively associated with audit fees.

2.5.2. Moderating effect of AEM

Although REM and AEM both increase audit risk, their detectability and prominence are different. Accrual error monitoring (AEM) provides a visual signal for auditors since it is quantitative, standardized, and integrated directly in financial statements (Abbott et al., 2006). On the other hand, REM is more difficult to separate since it is embedded in actual operational actions that frequently imitate valid business strategies (Choi et al., 2018). The notion of signals (Spence, 1978) states that auditors' attentional resources are monopolized by the dominant signal, which is AEM when it is pronounced. Due to time and mental limitations, auditors may put off operational studies that call for a more in-depth understanding of the business in favor of examining accrual abnormalities, for which detection mechanisms are already in place (Sohn, 2011; Yu, 2020). We explain how auditors prioritize competing risk signals by extending signaling theory, which helps to comprehend the moderating role of AEM. The auditors' interpretation and prioritization of the two earnings management strategies as competing risk signals is transformed by the stronger and more noticeable signal that AEM gives within this framework compared to REM. The empirical evidence backs up this method of prioritization: auditors pay greater attention to anomalies based on accruals and may pay less attention to the subtler signals linked with REM (Sohn, 2011; Yu, 2020). Consequently, when AEM is pronounced, it absorbs a disproportionate share of auditors' cognitive attention, reducing the incremental influence of REM on audit assessments and fee pricing. This mechanism provides the theoretical bridge that explains why AEM moderates the REM–audit fee relationship. This attentional bias implies that high levels of AEM may inadvertently mask or dilute the perceived severity of concurrent REM activities. This mechanism finds indirect support in recent literature. Similarly, Sitanggang et al. (2020) note that auditors' attention is conditional on the larger earnings management environment, since they see diverse fee reactions to various REM components. If auditors in Iran are too busy looking for anomalies in accruals to notice operational smoothing, then it stands to reason that smoother cash flow patterns, which could be an indication of REM, would be associated with reduced abnormal fees (Azadeh et al., 2023). Because of this theoretical uncertainty and the fact that previous research has not agreed on which way AEM acts as a moderator, we provide a non-directional hypothesis:

H2: AEM moderates the relationship between REM and audit fees.

This hypothesis advances audit pricing theory by integrating insights from behavioral decision-making into traditional risk models, offering a more realistic account of how auditors process multiple, concurrent risk signals in practice.

3. Research methodology

This applied, correlational study employs an ex post facto design, utilizing historical data. The study hypotheses are tested using panel regression models, which are a quantitative methodology. Statistical techniques were carried out in Stata (version 14), following initial data preprocessing in Excel. The theoretical framework was built through library research and documentary analysis. Sources for the data included audited financial accounts and official databases such as Rahavard Novin, the Statistical Center of Iran, Codal, and the Tehran Stock Exchange. The study population consists of companies that have been listed on the Tehran Stock Exchange from 2013 to 2025. These companies were chosen based on certain criteria: (1) they have been listed continuously since 2013; (2) they are not involved in financial, insurance, or banking activities; (3) their fiscal year ends in Esfand (March) and has not changed throughout the study period; (4) there must be at least eight firms per industry for effective cross-sectional earnings management estimation, in accordance with Banimahd et al. (2022) and Hajiha and Chenari Bouket (2016), and (5) information about audit fees must be publicly available (Banimahd et al. (2022) and Hajiha and Chenari Bouket (2016)). The research period (2013–2025) is reported in Gregorian years corresponding to the Iranian fiscal years 1391–1402 (2012-2023). Each fiscal year spans parts of two Gregorian years. Considering the study period and the aforementioned restrictions, the total number of observations is 864 firm-years. The data screening process is presented in Table 1.

Table 1. Final sample selection steps

Description	Number of Firms	Number of Observations
Total listed Companies by the end of 2025	460	5520
Subtracted: Companies not listed since 2013	(34)	(408)
Subtracted: Companies delisted from 2013 to 2025	(74)	(888)
Subtracted: Specialized investment, insurance, and bank companies	(73)	(876)
Subtracted: Companies with a fiscal year not end in Esfand	(76)	(912)
Subtracted: Companies that changed their fiscal year	(25)	(300)
Subtracted: Industries with fewer than eight companies	(35)	(420)
Subtracted: Companies with unavailable data or not disclosing audit fees	(71)	(852)
Remaining	72	864

3.1. Data analysis method and testing of classical regression assumptions

A multivariate panel regression was used to test the hypotheses put forth, and a series of diagnostic checks was performed to ensure compliance with classical regression assumptions. Variance inflation factor (VIF) values were all below ten, suggesting the absence of serious multicollinearity. Indeed, the panel structure of our dataset and the inclusion of industry and year controls led us to expect that heteroscedasticity and autocorrelation were present. To address both issues simultaneously, the models were estimated using ordinary least squares with robust standard errors (vce(robust)). To mitigate potential distortions from non-normality and outliers, variables were log-transformed, standardized, or trimmed as appropriate. The large sample size, in conjunction with the central limit theorem, helped minimize the impact of residual non-normality (Aflatooni, 2018). To control for cross-sectional heterogeneity, industry-specific structural variables were incorporated into the models. Instead of employing year fixed effects, annual inflation rates were used to capture temporal variation. By embedding these control variables directly into the model structure, the fixed effects of industry and time were effectively absorbed, rendering tests such as the F-Limer (for panel suitability) and Hausman (for model specification) unnecessary.

3.2. Research models

Model for testing the hypotheses

To test the research hypotheses, this study adopts the empirical framework proposed by Greiner et al. (2017), along with extended model specifications tailored to the Iranian research context by Mohammadrezaei and Faraji (2019). Model (1) is used to test the main relationship between REM and audit fees (H1), while Model (2) tests the moderating role of AEM in the REM–audit fee relationship (H2).

Model (1)

$$\begin{aligned} LNAUDIT\ FEES_{i,t} &= \beta_0 + \beta_1 REM_{i,t} + \beta_2 AEM_{i,t} + \beta_3 LNASSETS_{i,t} + \beta_4 \mathrm{INVRECV}_{i,t} + \beta_5 LEV_{i,t} \\ &+ \beta_6 \mathrm{EBIT}_{i,t} + \beta_7 LOSS_{i,t} + \beta_8 AUDITOR\ CHANGE_{i,t} + \beta_9 SUBS_{i,t} + \beta_{10} FRA_{i,t} \\ &+ \beta_{11} LNDELAY_{i,t} + \beta_{12} LNAGE_{i,t} + \beta_{13} DISTRESS_{i,t} + \beta_{14} CEO\ CHANGES_{i,t} \\ &+ \beta_{15} AUDIT\ OPINION_{i,t} + \beta_{16} \mathrm{INFLATION}_{i,t} + \beta_{17} \mathrm{DPS}_{i,t} + \beta_{18} \mathrm{AQPNum}_{i,t} \\ &+ \beta_{19} \mathrm{FCF}_{i,t} + \beta_{20} \mathrm{INS}_{i,t} + \beta_{21} \mathrm{CONOWN}_{i,t} + \beta_{22} \textstyle\sum IndustryDum_{i,t} + \varepsilon_{i,t} \end{aligned}$$

Model (2)

$$\begin{aligned} LNAUDIT\ FEES_{i,t} &= \beta_0 + \beta_1 REM_{i,t} + \beta_2 AEM_{i,t} + \beta_3 (REM * AEM)_{i,t} + \beta_4 LNASSETS_{i,t} \\ &+ \beta_5 \mathrm{INVRECV}_{i,t} + \beta_6 LEV_{i,t} + \beta_7 \mathrm{EBIT}_{i,t} + \beta_8 LOSS_{i,t} + \beta_9 AUDITOR\ CHANGE_{i,t} \\ &+ \beta_{10} SUBS_{i,t} + \beta_{11} FRA_{i,t} + \beta_{12} LNDELAY_{i,t} + \beta_{13} LNAGE_{i,t} + \beta_{14} DISTRESS_{i,t} \\ &+ \beta_{15} CEO\ CHANGES_{i,t} + \beta_{16} AUDIT\ OPINION_{i,t} + \beta_{17} \mathrm{INFLATION}_{i,t} + \beta_{18} \mathrm{DPS}_{i,t} \\ &+ \beta_{19} \mathrm{AQPNum}_{i,t} + \beta_{20} \mathrm{FCF}_{i,t} + \beta_{21} \mathrm{INS}_{i,t} + \beta_{22} \mathrm{CONOWN}_{i,t} \\ &+ \beta_{23} \textstyle\sum INDUSTRY\ DUM_{i,t} + \varepsilon_{i,t} \end{aligned}$$

3.3. Measurement of variables

3.2.1. Dependent variable

The dependent variable captures the outcome of interest in this study, audit pricing, as determined by external auditors.

LNAUDIT FEES = The natural logarithm of audit fees, which is extracted from the notes to the financial statements of the companies.

3.2.2. Independent variable

The primary independent variable reflects the extent to which firms engage in operational manipulation to manage reported earnings.

REM = REM index for the current year, calculated using the Roychowdhury (2006) model as detailed in Models 3, 4, and 5. To calculate the integrated REM measure, following Cohen and Zarowin (2010) and Zang (2012), the residuals of Models 3, 4, and 5 are summed together. Following their suggestion for simplicity in interpreting the resulting measure, the residuals in Models 3 and 5 are multiplied by -1. The higher this measure for a company, the higher the level of REM.

Primary REM Measurement Models (Roychowdhury, 2006):

Operating Cash Flow Model

Model (3)

$$\frac{CFO_t}{A_{t-1}} = \beta_0 + \beta_1 + \frac{1}{A_{t-1}} + \beta_2 \frac{SALES_t}{A_{t-1}} + \beta_3 \frac{\Delta SALES_t}{A_{t-1}} + \varepsilon$$

Cost of Goods Sold Model

Model (4)

$$\frac{CGS_t}{A_{t-1}} = \beta_0 + \beta_1 + \frac{1}{A_{t-1}} + \beta_2 \frac{SALES_t}{A_{t-1}} + \beta_3 \frac{\Delta SALES_t}{A_{t-1}} + \beta_4 \frac{\Delta SALES_{t-1}}{A_{t-1}} + \varepsilon$$

Discretionary Expenses Model

Model (5)

$$\frac{DISE_t}{A_{t-1}} = \beta_0 + \beta_1 + \frac{1}{A_{t-1}} + \beta_2 \frac{SALES_t}{A_{t-1}} + \varepsilon$$

Where:

CFO = Current year operating cash flow.

1/At−1= Used as a scaling factor to homogenize the model.

SALES = Total sales revenue for the current year.

ΔSALES = Change in sales revenue from the current year compared to the previous year.

CGS = Cost of goods sold for the current year.

ΔSALES (t-1) = Change in sales revenue from the previous year compared to two years ago.

DISE = General, distribution, and selling expenses for the current year.

To calculate the integrated measure of REM, following Cohen and Zarowin (2010) and Zang (2012), the residuals of Models 3, 4, and 5 are summed.

Alternative Approach to Measuring REM

To evaluate the robustness of the main findings, this study employed the discretionary expenses model as specified by Cohen and Zarowin (2010). This model extends the original framework introduced by Roychowdhury (2006). Cohen and Zarowin (2010) highlighted a limitation of the original model, noting that when managers artificially boost sales in a given year, the resulting residuals from the discretionary expense estimation may appear insignificant, thereby masking REM behavior. To overcome this limitation, they proposed incorporating lagged sales into the model, allowing discretionary expenses to be estimated as a function of prior-year sales rather than contemporaneous figures. Following their suggestion for simplicity in interpreting the resulting measure, the residuals in Models 6 and 8 are multiplied by -1. The higher this measure for a company, the higher the level of REM. This alternative specification is used solely for robustness testing and does not replace the primary REM measurement presented earlier. In line with this adjustment, we implemented the revised model to conduct robustness tests and determine whether our primary results hold under an alternative specification of REM. The second method for estimating REM includes the following models:

Operating Cash Flow Model

Model (6)

$$\frac{CFO_t}{A_{t-1}} = \beta_0 + \beta_1 + \frac{1}{A_{t-1}} + \beta_2 \frac{SALES_t}{A_{t-1}} + \beta_3 \frac{\Delta SALES_t}{A_{t-1}} + \varepsilon$$

Cost of Goods Sold Model

Model (7)

$$\frac{CGS_t}{A_{t-1}} = \beta_0 + \beta_1 + \frac{1}{A_{t-1}} + \beta_2 \frac{SALES_t}{A_{t-1}} + \beta_3 \frac{\Delta SALES_t}{A_{t-1}} + \beta_4 \frac{\Delta SALES_{t-1}}{A_{t-1}} + \varepsilon$$

Discretionary Expenses Model

Model (8)

$$\frac{DISE_t}{A_{t-1}} = \beta_0 + \beta_1 + \frac{1}{A_{t-1}} + \beta_2 \frac{SALES_{t-1}}{A_{t-1}} + \varepsilon$$

This model is similar to REM Models (3), (4), and (5), with the difference that in Model (8), sales changes are lagged by one period (i.e., the previous year's sales).

3.2.3. Control variables

Following prior audit pricing literature, a set of control variables is included to account for firm-specific characteristics, governance attributes, and macroeconomic conditions that may influence audit fees.

AEM = The AEM index for the current year, derived through the following models.

In the main analysis, AEM is measured using the Kothari et al. (2005) model; the Jones, Modified Jones, and Kasznik models are employed only in robustness tests to assess the sensitivity of the results to alternative AEM specifications.

All AEM measures retain their original sign to preserve directional information about earnings management.

AEM Models

Jones (*1991*) Model:

$$\frac{TACC_t}{A_{t-1}} = \beta_0 + \beta_1 + \frac{1}{A_{t-1}} + \beta_2 \frac{\Delta REV_t}{A_{t-1}} + \beta_3 \frac{PPE_t}{A_{t-1}} + \varepsilon$$

Model (9)

Where:

TACC = Total accruals (discretionary accruals).

$1/A_{t-1}$= Used as a scaling factor to homogenize the model.

ΔREV = Change in total revenue for the current year compared to the previous year.

PPE = Property, plant, and equipment.

ε = Model error term, indicating AEM.

To scale the variables, all variables are divided by total assets at the beginning of the period.

Data Arrangement for Residuals Calculation

To compute the residuals for the models, the data should be segmented by industry and year with a cross-sectional structure. Additionally, there must be at least eight companies in each industry.

Modified Jones Model (*Dechow et al., 1995*):

$$\frac{TACC_t}{A_{t-1}} = \beta_0 + \beta_1 + \frac{1}{A_{t-1}} + \beta_2 \frac{\Delta REV_t - \Delta REC_t}{A_{t-1}} + \beta_3 \frac{PPE_t}{A_{t-1}} + \varepsilon$$

Model (10)

Where:

ΔREC = Change in accounts receivable from the current year compared to the previous year. The remaining variables are the same as in the Jones model.

Kasznik (*1999*) Model:

$$\frac{TACC_t}{A_{t-1}} = \beta_0 + \beta_1 + \frac{1}{A_{t-1}} + \beta_2 \frac{\Delta REV_t - \Delta REC_t}{A_{t-1}} + \beta_3 \frac{PPE_t}{A_{t-1}} + \beta_4 + \frac{\Delta CFO_t}{A_{t-1}} + \varepsilon$$

Model (11)

Where:

ΔCFO = Change in operating cash flow from the current year compared to the previous year. The remaining variables are the same as in the modified Jones model.

Kothari et al. (*2005*) Model:

$$\frac{TACC_t}{A_{t-1}} = \beta_0 + \beta_1 + \frac{1}{A_{t-1}} + \beta_2 \frac{\Delta REV_t - \Delta REC_t}{A_{t-1}} + \beta_3 \frac{PPE_t}{A_{t-1}} + \beta_4 + ROA_t + \varepsilon$$

Model (12)

Where:

ROA = Return on assets, calculated as net income divided by total assets. The remaining variables are the same as in the Kasznik model.

LNASSETS = The natural logarithm of total assets, representing company size.

INVRECV = The ratio of accounts receivable and inventory to total assets.

LEV = The ratio of total liabilities to total assets.

EBIT = The ratio of earnings before interest and taxes to total assets.

LOSS = A binary variable that takes the value of 1 if the company reported a loss in the current year and 0 otherwise.

AUDITOR CHANGE = A binary variable that takes the value of 1 if the company changed its auditor, and 0 otherwise.

SUBS = A binary variable indicating company complexity, taking the value of 1 if the company has subsidiaries and 0 otherwise (Greiner et al., 2017; Mohammadrezaei and Faraji, 2019).

FRA = First-tier auditors, with a value of 1 if the auditing firm is a first-tier institution trusted by the Securities and Exchange Organization, and 0 otherwise (Mohammadrezaei and Faraji, 2019).

LNDELAY = Audit report delay, calculated as the natural logarithm of the number of days between the end of the fiscal year and the audit report date (Greiner et al., 2017).

LNAGE = The natural logarithm of the company's age, calculated from the founding year to the reporting period (Mohammadrezaei and Faraji, 2019).

DISTRESS = A binary variable indicating bankruptcy, with a value of 1 if the company is bankrupt and 0 if it is not. This indicator is derived using the Z-Score model (Zmijewski, 1984) as follows:

$$Z = -4.3 - 4.5(ROA) + 5.7(FINL) + 0.004(LIQ)$$

Model (13)

Where:

Z = Bankruptcy probability. If Z < 0.5, the company is considered non-bankrupt, and if Z > 0.5, the company is considered bankrupt.

ROA = Return on assets (net income divided by total assets).

FINL = Total liabilities to total assets ratio.

LIQ = Current assets to current liabilities ratio (current ratio).

AUDIT OPINION = The audit opinion, with a value of 1 if the auditor provides an adverse opinion and 0 otherwise.

CEO CHANGES = A binary variable indicating CEO change, with a value of 1 if the CEO changed, and 0 otherwise.

INFLATION = The annual inflation rate, extracted from the website of the Statistical Center of Iran.

DPS = The amount of dividends paid by the company, standardized by share price.

AQPNum = The number of qualified opinion paragraphs in the auditor's report. (Bani Mahd et al., 2014).

FCF = Free cash flow, calculated as the difference between operating cash flow and net investment in fixed assets. This variable is standardized by total assets.

INS = The percentage of institutional shareholders, derived from the notes accompanying the companies' financial statements.

CONOWN = Ownership concentration, defined as shareholders owning more than 50% of the company's shares. This information is extracted from the notes to the companies' audited financial statements.

INDUSTRYDUM = Industry control variable.

3.3. Moderator variable

To test the core theoretical proposition regarding dual earnings management, we include a moderating variable that captures the interaction between real and accrual-based strategies.

REM × AEM = The interaction between REM and AEM, moderating the relationship between REM and audit fees.

4. Findings and discussion

4.1. Descriptive statistics

Table 2 presents the descriptive statistics of the study variables, summarizing key measures of central tendency, dispersion, and distributional shape, including the mean, median, standard deviation, skewness, and kurtosis.

Tables 2 and 3 present the descriptive statistics for the main variables used in the analysis. The mean value of the natural logarithm of audit fees is 6.949 (median = 6.870), with a standard deviation of 0.908, suggesting moderate dispersion across firms. Both REM and AEM exhibit mean values close to zero, consistent with their residual-based estimation methods. Among the control variables, firm size has an average logarithmic value of 14.032, while inventories and receivables comprise 49.6% of total assets, reflecting substantial investment in working capital. The average leverage ratio is 58.5%, and EBIT accounts for 14.6% of total assets, implying a recovery period of approximately seven years. The mean value of audit delay is 4.267 (logarithmic days), and the average firm age is 3.537. Dividend payout and free cash flow are relatively low, with mean values of 0.076 and 0.079, respectively. On average, firms receive 1.39 audit qualification paragraphs, and institutional investors hold 74.3% of outstanding shares. Ownership concentration averages 53.8%, while the annual inflation rate across the study period averages 25.9%, reflecting substantial macroeconomic volatility. Regarding binary variables, 48.0% of firms receive modified audit opinions, and 68.9% are audited by first-tier audit firms. Additionally, 11.8% of firms report negative earnings, 23.0% operate subsidiaries, 4.9% are classified as financially distressed, 23.2% change their external auditor annually, and 31.6% experience CEO turnover.

Table 2. Descriptive statistics of continuous variables

Symbol	Mean	Median	Maximum	Minimum	Standard deviation	Skewness	Kurtosis
LNAUDIT FEES	6.949	6.870	8.730	5.440	0.908	0.235	2.263
REMCFO	0.000	0.004	0.302	-0.396	0.092	-0.6747	2.559
REMCGS	0.000	0.000	0.335	-0.563	0.083	-0.195	3.152
REMDISE	0.000	0.000	0.088	-0.077	0.024	0.163	2.465
REM1	0.000	0.004	0.662	-0.795	0.156	-0.878	3.011
REM2	0.000	0.004	0.654	-0.807	0.156	-0.912	2.045
AEM_{JONES}	0.000	-0.001	0.318	-0.265	0.078	0.357	2.892
$AEM_{MODIFIED}$ JONES	0.000	0.000	0.336	-0.316	0.082	0.251	3.193
$AEM_{KASZNIK}$	0.000	0.000	0.311	-0.281	0.063	0.195	2.991
$AEM_{KOTHARI}$	0.000	0.000	0.335	-0.324	0.063	0.058	2.456
LNASSETS	14.032	13.920	16.650	11.870	1.246	0.306	2.499
INVRECV	0.496	0.500	0.820	0.130	0.208	-0.136	2.937
LEV	0.585	0.620	0.920	0.170	0.202	-0.351	2.383
EBIT*	0.146	0.120	0.460	-0.080	0.144	0.611	2.686
LNDELAY	4.267	4.330	4.740	3.580	0.364	-0.404	2.866
LNAGE	3.537	3.625	4.090	2.710	0.430	-0.476	2.976
DPS**	0.076	0.060	0.230	0.000	0.076	0.609	2.061
AQPNum	1.390	0.000	15.000	0.000	2.060	0.089	2.815
FCF*	0.079	0.070	0.320	-0.120	0.113	0.372	2.642
INS	0.743	0.780	0.950	0.230	0.186	-0.213	3.003
CONOWN	0.538	0.520	0.880	0.170	0.187	0.017	2.562
INFLATION	0.259	0.276	0.458	0.070	0.129	0.061	3.361

* It has been scaled by total assets.
** It has been scaled by stock price.

Table 3. Descriptive statistics of discrete variables

Symbol	Frequency		Percentage	
	1	0	1	0
AUDIT OPINION	415	449	48.030	51.970
FRA	595	269	68.870	31.130
LOSS	102	762	11.810	88.190
AUDITOR CHANGE	200	664	23.150	76.850
SUBS	199	665	23.030	76.970
DISTRESS	822	42	95.140	4.860
CEO CHANGES	591	273	68.400	31.600

Although formal normality tests such as the Jarque–Bera test were not conducted, this omission is methodologically justified given the large sample size (N = 864). According to the Central Limit Theorem, the sampling distribution of regression estimators approaches normality as sample size increases, even when the underlying variables exhibit non-normality (Aflatooni, 2018). This justifies the use of parametric inference without requiring strict normality of the raw variables.

In addition, Table 2 reports skewness and kurtosis values to assess the distributional shape and normality of the variables. The results showed that most variables exhibit moderate skewness and kurtosis values, generally within acceptable statistical ranges, suggesting that the data are approximately normally distributed. Given the large sample size and the use of robust standard errors, the analytical framework remains both statistically valid and methodologically robust.

4.2. Hypothesis testing results

4.2.1. Hypothesis one test result

Table 4 presents the regression results testing Hypothesis 1, which examines the positive association between REM and audit fees.

Table 4. Regression results for the relationship between REM and audit fees (dependent variable: audit fees)

Symbol	Coefficient	t-Statistic
REM1	0.580**	3.240
$AEM_{KOTHARI}$	0.517**	2.580
LNASSETS	0.605**	20.630
INVRECV	0.360*	2.430
LEV	0.240	1.200
EBIT	0.854**	2.950
LOSS	0.152*	1.780
AUDITOR CHANGE	-0.110*	-1.850
SUBS	-0.080	-0.800
FRA	0.224**	3.780
LNDELAY	0.253**	2.680
LNAGE	0.381**	5.220
DISTRESS	0.053	0.380
CEO CHANGES	-0.068	-1.150
AUDIT OPINION	-0.003	-0.020
INFLATION	-0.364	-1.160
DPS	-2.144**	-5.280
AQPNum	-0.103**	-3.880
FCF	0.263	0.560
INS	0.486**	2.300
CONOWN	-0.193	-1.330
C	-2.634**	-4.880
IND	Controlled	
F-statistic	68.341	
Significance of the F-statistic	0.000	
R-squared	0.630	
Number of Observations	864	

** Significant at the 95% confidence level, * significant at the 90% confidence level

As shown in Table 4, REM exhibits a positive and statistically significant effect on audit fees (coefficient = 0.580; t = 3.24), indicating that auditors charge higher fees when firms engage more in REM activities, consistent with the notion that REM increases audit risk and effort. AEM, measured here as discretionary accruals, also positively influences audit fees (coefficient = 0.517; t = 2.58), supporting its direct role in increasing audit complexity and cost. Firm size (LNASSETS), a key control variable, shows a strong positive association with audit fees (coefficient = 0.605; t = 20.63), reflecting the increased audit scope and resources required for larger firms. The ratio of inventory and receivables to total assets (INVRECV) is positively related to audit fees (coefficient = 0.360; t = 2.43), likely due to the higher inherent risk in verifying these assets. Leverage (LEV) has a positive but not statistically significant effect (coefficient = 0.240; t = 1.20), while the EBIT ratio (EBIT) is positively and significantly associated with audit fees (coefficient = 0.854; t = 2.95), suggesting that more profitable firms might incur higher audit costs. Loss presence (LOSS) also shows a positive and marginally significant impact (coefficient = 0.152; t = 1.78). Auditor change (CHANGE) is negatively and significantly related to audit fees (coefficient = –0.110; t = –1.85), indicating that a recent auditor switch might reduce fees, possibly due to introductory pricing or reduced risk assessments. Other controls, such as first-rank auditor status (FRA), audit report delay (LNDELAY), and firm age (LNAGE), show positive and significant coefficients, reaffirming their role in increasing audit fees. In contrast, dividend per share (DPS) and the number of qualified opinion paragraphs (AQPNum) negatively and significantly affect audit fees, possibly reflecting less perceived audit risk or negotiation pressure on fees. Variables such as subsidiary presence (SUBS), financial distress (DISTRESS), CEO changes, audit opinion, inflation, free cash flow, and ownership concentration do not exhibit statistically significant relationships. The model explains a substantial portion of the variation in audit fees, with an R-squared of 0.630. The overall model fit is strong, supported by a highly significant F-statistic of 68.341 ($p < 0.001$), confirming the robustness and reliability of the findings.

4.2.2. Hypothesis two test result

The second hypothesis of this study investigates whether AEM moderates the relationship between REM and audit fees. Table 5 reports the regression results for the corresponding model used to test this interaction effect.

Table 5. Regression results for the moderating effect of AEM on the REM–audit fee relationship (dependent variable: audit fees)

Symbol	Coefficient	t-Statistic
REM1	0.572**	3.100
$AEM_{KOTHARI}$	0.308**	2.850
Interaction	-2.640**	-2.820
LNASSETS	0.508**	22.820
INVRECV	0.241*	1.850
LEV	0.264	1.620
EBIT	0.802**	2.840
LOSS	0.154*	1.840
AUDITOR CHANGE	-0.095*	-1.920
SUBS	-0.074	-1.030
FRA	0.207**	4.410
LNDELAY	0.217**	2.630
LNAGE	0.346**	6.170
DISTRESS	0.042	0.350
CEO CHANGES	-0.052	-1.100
AUDIT OPINION	-0.007	-0.090
INFLATION	-0.301	-1.270
DPS	-2.113**	-6.050
AQPNum	-0.098**	-3.810
FCF	0.286	1.040
INS	0.474**	3.180
CONOWN	-0.189	-1.400
C	-2.570**	-4.850
IND	Controlled	
F-statistic	47.400	
Significance of the F-statistic	0.000	
R-squared	0.616	
Number of Observations	864	

** Significant at the 95% confidence level, * significant at the 90% confidence level

As shown in Table 5, the interaction term between REM and AEM, denoted as $REM \times AEM_{KOTHARI}$, is negative and statistically significant at the 95% confidence level (coefficient = –2.640; t = –2.82). This finding indicates that higher levels of AEM weaken the positive association between REM and audit fees, thereby supporting the hypothesis that AEM moderates this relationship. A marginal effect analysis further confirms this moderating pattern. When AEM is at one standard deviation below its mean, the marginal effect of REM on audit fees remains strongly positive and significant. However, as AEM increases to one standard deviation above its mean, this marginal effect diminishes substantially and may become statistically insignificant. This pattern suggests that at high levels of accrual manipulation, auditors focus less on real activity deviations, consistent with a dampening (rather than amplifying) moderation mechanism. Moreover, the coefficient on REM is positive and significant (coefficient = 0.572; $p < 0.05$), suggesting that auditors increase their fees in response to heightened REM activity, consistent with the view that REM increases audit risk. AEM, as measured by the Kothari model, also exhibits a positive and statistically significant effect on audit fees (coefficient = 0.308), confirming its direct role in driving increases in audit costs. Among the control variables, firm size, EBIT, firm age, and the presence of a first-tier auditor are all positively associated with audit fees, aligning with prior literature that links these factors to increased audit effort and complexity. In contrast, dividend per share and the number of audit qualification paragraphs show negative and significant relationships with audit fees, potentially reflecting reduced audit risk or downward pressure on audit pricing. The model's explanatory power is robust, as indicated by an R-squared value of 0.616. The F-statistic of 47.400 (p < 0.001) further confirms the model's overall significance and reliability.

4.3. Robustness tests of the research hypothesis

4.3.1. First robustness test

To evaluate the robustness of the research hypothesis, a series of sensitivity analyses was performed using four alternative models for measuring AEM. In these robustness tests, REM was consistently estimated using operating cash flows, while AEM was calculated based on the Jones, Modified Jones, Kasznik, and Kothari models. The results of these analyses are summarized in Tables 6 to 10.

Table 6. Robustness test results using operating cash flows to estimate REM (dependent variable: audit fees)

Symbol	AEM_JONES	AEM_MODIFIED JONES	AEM_KASZNIK	AEM_KOTHARI
REMCFO	1.181** (3.770)	1.181** (3.840)	1.035** (3.540)	1.175** (3.670)
AEMJONES	0.671** (2.100)	-	-	-
AEMMODIFIED JONES	-	0.722** (2.400)	-	-
AEMKASZNIK	-	-	0.724** (2.140)	-
AEMKOTHARI	-	-	-	0.655* (1.710)
REMCFO*AEMJONES	-0.725 (-0.430)	-	-	-
REMCFO*AEMMODIFIED JONES	-	0.247 (0.150)	-	-
REMCFO*AEMKASZNIK	-	-	-0.528 (-0.220)	-
REMCFO*AEMKOTHARI	-	-	-	-2.731** (-2.390)
LNASSETS	0.507** (22.880)	0.506** (22.750)	0.506** (22.860)	0.51** (22.920)
INVRECV	0.235* (1.800)	0.237* (1.810)	0.251* (1.930)	0.228* (1.740)
LEV	0.239 (1.440)	0.242 (1.480)	0.258 (1.570)	0.291* (1.790)
EBIT	0.788** (2.750)	0.826** (2.850)	0.779** (2.730)	0.739** (2.670)
LOSS	0.141* (1.680)	0.145* (1.720)	0.149* (1.750)	0.146* (1.760)
AUDITOR CHANGE	-0.093* (-1.880)	-0.09* (-1.820)	-0.095* (-1.910)	-0.094* (-1.900)
SUBS	-0.070 (-0.980)	-0.066 (-0.920)	-0.061 (-0.860)	-0.077 (-1.080)
FRA	0.214** (4.610)	0.217** (4.660)	0.217** (4.680)	0.215** (4.610)
LNDELAY	0.226** (2.770)	0.227** (2.790)	0.228** (2.800)	0.214** (2.610)
LNAGE	0.331** (5.910)	0.326** (5.850)	0.337** (5.980)	0.336** (5.960)
DISTRESS	0.047 (0.390)	0.053 (0.440)	0.058 (0.470)	0.040 (0.330)
CEO CHANGE	-0.056 (-1.190)	-0.051 (-1.070)	-0.051 (-1.090)	-0.050 (-1.050)
AUDIT OPINION	0.001 (0.010)	-0.002 (-0.030)	0.001 (0.010)	-0.009 (-0.120)
INFLATION	-0.313 (-1.310)	-0.332 (-1.390)	-0.309 (-1.300)	-0.287 (-1.210)
DPS	-2.152** (-6.200)	-2.156** (-6.190)	-2.163** (-6.190)	-2.144** (-6.110)
AQPNum	-0.101** (-3.880)	-0.1** (-3.890)	-0.101** (-3.920)	-0.098** (-3.790)
FCF	0.379 (1.330)	0.352 (1.210)	0.436 (1.510)	0.443 (1.550)
INS	0.434** (2.910)	0.439** (2.960)	0.429** (2.860)	0.429** (2.900)
CONOWN	-0.191 (-1.420)	-0.189 (-1.400)	-0.175 (-1.290)	-0.178 (-1.310)
C	-2.518** (-4.780)	-2.52** (-4.800)	-2.568** (-4.840)	-2.535** (-4.760)
IND	Controlled	Controlled	Controlled	Controlled
F-statistic	46.610	46.600	46.980	47.860
Significance of the F-statistic	0.000	0.000	0.000	0.000
R-squared	0.616	0.617	0.616	0.616
Number of Observations	864	864	864	864

Table 7. Robustness test results using cost of goods sold to estimate REM (dependent variable: audit fees)

Symbol	**AEM_JONES**	**AEM_MODIFIED JONES**	**AEM_KASZNIK**	**AEM_KOTHARI**
REMCGS	0.881** (2.400)	0.846** (2.390)	0.789** (2.150)	0.888** (2.670)
AEMJONES	**0.019 (2.060)		-	-
AEMMODIFIED JONES	-	0.166** (2.540)	-	-
AEMKASZNIK	-	-	0.305** (2.810)	-
AEMKOTHARI	-	-	-	0.012** (2.030)
REMCGS*AEMJONES	0.357 (0.190)	-	-	-
REMCGS*AEMMODIFIED JONES	-	0.371 (0.170)	-	-
REMCGS*AEMKASZNIK	-	-	1.325 (0.680)	-
REMCGS*AEMKOTHARI	-	-		-6.798** (-2.070)
LNASSETS	0.504** (22.730)	0.504** (22.760)	0.504** (22.710)	0.507** (22.740)
INVRECV	0.281** (2.200)	0.283** (2.210)	0.287** (2.230)	0.264** (2.050)
LEV	0.215 (1.300)	0.207 (1.260)	0.203 (1.220)	0.215 (1.310)
EBIT	0.752** (2.630)	0.776** (2.700)	0.782** (2.730)	0.784** (2.760)
LOSS	0.154* (1.830)	0.152* (1.810)	0.155* (1.840)	0.163* (1.930)
AUDITOR CHANGE	-0.093* (-1.880)	-0.092* (-1.850)	-0.091* (-1.840)	-0.097* (-1.960)
SUBS	-0.049 (-0.690)	-0.048 (-0.670)	-0.044 (-0.620)	-0.065 (-0.900)
FRA	0.217** (4.650)	0.217** (4.610)	0.217** (4.630)	0.208** (4.450)
LNDELAY	0.216** (2.620)	0.216** (2.620)	0.22** (2.670)	0.217** (2.650)
LNAGE	0.34** (6.050)	0.337** (6.030)	0.339** (6.040)	0.342** (6.120)
DISTRESS	0.037 (0.300)	0.041 (0.340)	0.039 (0.320)	0.056 (0.470)
CEO CHANGES	-0.055 (-1.170)	-0.054 (-1.140)	-0.054 (-1.150)	-0.053 (-1.130)
AUDIT OPINION	0.003 (0.040)	0.002 (0.020)	0.004 (0.060)	-0.005 (-0.060)
INFLATION	-0.287 (-1.210)	-0.295 (-1.240)	-0.292 (-1.230)	-0.308 (-1.300)
DPS	-2.108** (-6.120)	-2.086** (-6.080)	-2.088** (-6.120)	-2.1** (-6.170)
AQPNum	-0.102** (-3.940)	-0.102** (-3.960)	-0.104** (-4.010)	-0.1** (-3.870)
FCF	0.144 (0.530)	0.067 (0.240)	0.071 (0.290)	0.117 (0.450)
INS	0.473** (3.140)	0.468** (3.110)	0.465** (3.080)	0.487** (3.270)
CONOWN	-0.180 (-1.330)	-0.178 (-1.330)	-0.173 (-1.290)	-0.184 (-1.380)
C	-2.487** (-4.740)	-2.476** (-4.730)	-2.493** (-4.750)	-2.538** (-4.860)
IND	Controlled	Controlled	Controlled	Controlled
F-statistic	47.270	47.330	47.290	48.180
Significance of the F-statistic	0.000	0.000	0.000	0.000
R-squared	0.613	0.614	0.614	0.616
Number of Observations	864	864	864	864

Table 8. Robustness test results using discretionary expenses to estimate REM (dependent variable: audit fees)

Symbol	AEM_JONES	AEM_MODIFIED JONES	AEM_KASZNIK	AEM_KOTHARI
REMDISE	1.412** (2.530)	1.428** (2.560)	1.387** (2.520)	1.515** (2.640)
AEMJONES	0.241** (2.800)	-	-	-
AEMMODIFIED JONES	-	0.340** (2.180)	-	-
AEMKASZNIK	-	-	0.577** (2.720)	-
AEMKOTHARI	-	-	-	0.029** (2.080)
REMDISE*AEMJONES	-7.228 (-0.660)	-	-	-
REMDISE*AEMMODIFIED JONES	-	-14.110 (-1.480)	-	-
REMDISE*AEMKASZNIK	-	-	-12.865 (-1.230)	-
REMDISE*AEMKOTHARI	-	-	-	-19.704** (-2.650)
LNASSETS	0.509** (23.020)	0.509** (23.040)	0.508** (23.040)	0.51** (23.080)
INVRECV	0.328** (2.560)	0.322** (2.510)	0.325** (2.550)	0.316** (2.490)
LEV	0.228 (1.380)	0.226 (1.390)	0.224 (1.380)	0.223 (1.380)
EBIT	0.540* (1.910)	0.564** (1.990)	0.582** (2.060)	0.486* (1.760)
LOSS	0.135 (1.600)	0.132 (1.570)	0.138 (1.620)	0.129 (1.530)
AUDITOR CHANGE	-0.1** (-2.020)	-0.099** (-2.000)	-0.1** (-2.020)	-0.1** (-2.020)
SUBS	-0.047 (-0.650)	-0.047 (-0.650)	-0.042 (-0.580)	-0.045 (-0.630)
FRA	0.225** (4.850)	0.227** (4.880)	0.223** (4.790)	0.23** (4.940)
LNDELAY	0.211** (2.570)	0.206** (2.520)	0.206** (2.520)	0.205** (2.490)
LNAGE	0.315** (5.430)	0.312** (5.410)	0.319** (5.530)	0.32** (5.520)
DISTRESS	0.050 (0.440)	0.064 (0.560)	0.072 (0.600)	0.035 (0.310)
CEO CHANGES	-0.048 (-1.010)	-0.044 (-0.940)	-0.046 (-0.980)	-0.050 (-1.060)
AUDIT OPINION	0.003 (0.040)	0.001 (0.010)	0.002 (0.020)	0.002 (0.020)
INFLATION	-0.220 (-0.930)	-0.231 (-0.970)	-0.226 (-0.950)	-0.219 (-0.930)
DPS	-2.001** (-5.880)	-2.003** (-5.870)	-2.003** (-5.880)	-2.1** (-6.170)
AQPNum	-0.107** (-4.090)	-0.106** (-4.060)	-0.107** (-4.080)	-0.105** (-4.020)
FCF	-0.042 (-0.160)	-0.074 (-0.280)	-0.042 (-0.170)	0.046 (0.180)
INS	0.389** (2.640)	0.383** (2.620)	0.382** (2.610)	0.386** (2.620)
CONOWN	-0.156 (-1.330)	-0.150 (-1.120)	-0.148 (-1.090)	-0.139 (-1.010)
C	-2.4** (-4.570)	-2.389** (-4.560)	-2.405** (-4.550)	-2.384** (-4.520)
IND	Controlled	Controlled	Controlled	Controlled
F-statistic	50.560	51.020	51.220	50.330
Significance of the F-statistic	0.000	0.000	0.000	0.000
R-squared	0.610	0.611	0.612	0.611
Number of Observations	864	864	864	864

Table 9. Robustness test results using composite REM (dependent variable: audit fees)

Symbol	AEMJONES	AEMMODIFIED JONES	AEMKASZNIK	AEMKOTHARI
REM1	0.554** (3.060)	0.559** (3.090)	0.521** (2.840)	0.572** (3.100)
AEMJONES	0.298** (2.050)			
AEMMODIFIED JONES		0.414** (2.440)		
AEMKASZNIK			0.482** (2.390)	
AEMKOTHARI				0.308** (2.850)
REM*AEMJONES	-0.177 (-0.170)			
REM*AEMMODIFIED JONES		0.269 (0.240)		
REM*AEMKASZNIK			0.477 (0.390)	
REM*AEMKOTHARI				-2.64** (-2.820)
LNASSETS	0.505** (22.840)	0.504** (22.750)	0.503** (22.770)	0.508** (22.820)
INVRECV	0.258** (1.990)	0.259** (2.000)	0.266** (2.050)	0.241* (1.850)
LEV	0.232 (1.400)	0.226 (1.390)	0.230 (1.400)	0.264 (1.620)
EBIT	0.797** (2.750)	0.832** (2.850)	0.816** (2.820)	0.802** (2.840)
LOSS	0.148* (1.760)	0.149* (1.780)	0.154* (1.820)	0.154* (1.840)
AUDITOR CHANGE	-0.092* (-1.850)	-0.09* (-1.810)	-0.091* (-1.830)	-0.095* (-1.920)
SUBS	-0.061 (-0.850)	-0.057 (-0.800)	-0.054 (-0.770)	-0.074 (-1.030)
FRA	0.214** (4.570)	0.215** (4.580)	0.215** (4.590)	0.207** (4.410)
LNDELAY	0.225** (2.720)	0.226** (2.740)	0.229** (2.780)	0.217** (2.630)
LNAGE	0.343** (6.120)	0.34** (6.070)	0.345** (6.130)	0.346** (6.170)
DISTRESS	0.037 (0.310)	0.041 (0.340)	0.043 (0.350)	0.042 (0.350)
CEO CHANGES	-0.057 (-1.200)	-0.054 (-1.140)	-0.055 (-1.160)	-0.052 (-1.100)
AUDIT OPINION	0.002 (0.020)	0.000 (0.000)	0.003 (0.040)	-0.007 (-0.090)
INFLATION	-0.311 (-1.300)	-0.324 (-1.360)	-0.315 (-1.320)	-0.301 (-1.270)
DPS	-2.145** (-6.190)	-2.138** (-6.160)	-2.154** (-6.210)	-2.113** (-6.050)
AQPNum	-0.10** (-3.860)	-0.101** (-3.900)	-0.102** (-3.930)	-0.098** (-3.810)
FCF	0.264 (0.940)	0.218 (0.760)	0.276 (1.010)	0.286 (1.040)
INS	0.472** (3.150)	0.473** (3.160)	0.469** (3.110)	0.474** (3.180)
CONOWN	-0.189 (-1.400)	-0.188 (-1.400)	-0.181 (-1.340)	-0.189 (-1.400)
C	-2.54** (-4.810)	-2.535** (-4.820)	-2.405** (-4.550)	-2.384** (-4.520)
IND	Controlled	Controlled	Controlled	Controlled
F-statistic	50.560	51.020	51.220	50.330
Significance of the F-statistic	0.000	0.000	0.000	0.000
R-squared	0.610	0.611	0.612	0.611
Number of Observations	864	864	864	864

Table 10. Robustness test results using the alternative REM model based on lagged sales (dependent variable: audit fees)

Symbol	AEMJONES	AEMMODIFIED JONES	AEMKASZNIK	AEMKOTHARI
REM2	0.531** (2.890)	0.534** (2.920)	0.495** (2.660)	0.548** (2.940)
AEM_JONES	0.296** (2.010)			
AEMMODIFIED JONES		0.405** (2.410)		
AEMKASZNIK			0.480** (2.390)	
AEMKOTHARI				0.295** (2.810)
REM2*AEMJONES	-0.177 (-0.180)			
REM2*AEMMODIFIED JONES		0.235 (0.210)		
REM2*AEMKASZNIK			0.494 (0.400)	
REM2*AEMKOTHARI				-2.753** (-2.860)
LNASSETS	0.505** (22.850)	0.504** (22.760)	0.504** (22.770)	0.508** (22.820)
INVRECV	0.258** (2.000)	0.260** (2.000)	0.267** (2.060)	0.241* (1.850)

LEV	0.236 (1.420)	0.231 (1.410)	0.234 (1.420)	0.269* (1.660)
EBIT	0.793** (2.740)	0.826** (2.830)	0.811** (2.800)	0.799** (2.820)
LOSS	0.147* (1.760)	0.149* (1.770)	0.153* (1.810)	0.153* (1.840)
AUDITOR CHANGE	-0.093* (-1.880)	-0.091* (-1.830)	-0.092* (-1.850)	-0.097* (-1.950)
SUBS	-0.060 (-0.850)	-0.057 (-0.800)	-0.054 (-0.760)	-0.074 (-1.030)
FRA	0.213** (4.560)	0.214** (4.570)	0.214** (4.580)	0.206** (4.390)
LNDELAY	0.224** (2.710)	0.226** (2.730)	0.229** (2.770)	0.217** (2.630)
LNAGE	0.343** (6.110)	0.34** (6.070)	0.345** (6.120)	0.346** (6.160)
DISTRESS	0.037 (0.300)	0.040 (0.330)	0.043 (0.350)	0.041 (0.340)
CEO CHANGES	-0.056 (-1.180)	-0.053 (-1.120)	-0.054 (-1.140)	-0.051 (-1.080)
AUDIT OPINION	0.003 (0.030)	0.001 (0.010)	0.004 (0.050)	-0.005 (-0.070)
INFLATION	-0.309 (-1.290)	-0.321 (-1.340)	-0.312 (-1.310)	-0.300 (-1.260)
DPS	-2.137** (-6.170)	-2.130** (-6.130)	-2.146** (-6.180)	-2.103** (-6.010)
AQPNum	-0.101** (-3.880)	-0.101** (-3.910)	-0.102** (-3.950)	-0.099** (-3.830)
FCF	0.254 (0.940)	0.210 (0.730)	0.264 (0.960)	0.279 (1.020)
INS	0.470** (3.130)	0.470** (3.140)	0.466** (3.090)	0.471** (3.160)
CONOWN	-0.188 (-1.390)	-0.186 (-1.380)	-0.179 (-1.330)	-0.188 (-1.400)
C	-2.542** (-4.820)	-2.537** (-4.820)	-2.568** (-4.850)	-2.574** (-4.860)
IND	Controlled	Controlled	Controlled	Controlled
F-statistic	46.550	46.390	46.600	47.540
Significance of the F-statistic	0.000	0.000	0.000	0.000
R-squared	0.614	0.615	0.615	0.616
Number of Observations	864	864	864	864

Note: Each cell presents the coefficient followed by its t-statistic in parentheses.
Statistical significance is indicated as follows:
** Significant at the 95% confidence level, * significant at the 90% confidence level

Robustness tests employing five alternative specifications of REM, operating cash flow (Table 6), cost of goods sold (Table 7), discretionary expenses (Table 8), composite REM model (Table 9), and the alternative REM model based on lagged sales (Table 10), consistently reinforce the study's primary findings. Across all models, REM maintains a positive and statistically significant association with audit fees, indicating that auditors perceive it as a major risk factor that necessitates increased audit effort. Similarly, AEM, measured using the Jones, Modified Jones, Kasznik, and Kothari models, exhibits a consistently positive and significant effect on audit fees, suggesting its independent contribution to perceived audit risk and cost. Notably, a significant negative interaction between REM and AEM is observed only when AEM is measured using the Kothari model. This finding implies that AEM, when estimated with higher precision, attenuates the positive relationship between REM and audit fees. No such moderating effect emerges under the Jones-type models, emphasizing the critical role of model selection in detecting dual earnings management interactions. Among the control variables, firm size, profitability (EBIT), firm age, audit delay, affiliation with a top-tier auditor, and institutional ownership all show robust positive effects on audit fees, consistent with expectations regarding audit complexity and risk. In contrast, dividend per share, auditor change, and the number of audit qualification paragraphs are negatively associated with audit fees, possibly reflecting lower perceived risk or pricing pressure. Other variables, including CEO turnover, financial distress, free cash flow, and ownership concentration, do not exhibit a statistically significant relationship with audit fees. Collectively, these robustness checks confirm the stability of the main results and underscore the importance of model precision in dual earnings management research.

5. Discussion

This study investigated whether AEM moderates the relationship between REM and audit fees. Regarding H1, the coefficient of REM (0.580, t = 3.24) reported in Table 4 indicates a positive and statistically significant relationship between REM and audit fees. This result suggests that auditors interpret higher levels of REM as a signal of increased audit risk, prompting additional audit effort and consequently higher fees. Firms engaging more intensively in operational manipulation thus face greater monitoring costs. This finding is consistent with the predictions of agency theory (Meckling and Jensen, 1976) and aligns with empirical evidence provided by Choi et al. (2018), Greiner et al. (2017), and Moradi et al. (2019), which collectively confirm that auditors incorporate the risk implications of REM into their pricing decisions. The interaction coefficient of –2.640 (t = –2.82) suggests that when firms simultaneously engage in both REM and AEM, commonly referred to as dual earnings management, the positive association between REM and audit fees weakens. This attenuation may result from auditors' diminished ability to isolate the primary source of reporting risk, thereby reducing the incremental effect of REM on audit pricing. This finding provides empirical support for the moderating role of AEM, which aligns with signaling theory (Spence, 1978) and posits that AEM may divert auditor attention away from operational manipulation due to its adverse informational signals. The result is consistent with prior studies by Alhadab (2018), who found that greater AEM intensity reduces auditors' sensitivity to REM, likely because accrual manipulation is perceived as a more prominent risk factor. Nevertheless, REM independently maintains a positive and statistically significant relationship with audit fees (coefficient = 0.572; t = 3.10), reinforcing the notion that auditors associate REM with elevated audit risk that necessitates increased effort and compensation. This outcome supports the predictions of agency theory (Meckling and Jensen, 1976) and aligns with the findings of Choi et al. (2018) and Greiner et al. (2017). In addition, AEM, measured using the Kothari model, demonstrates a positive and significant effect on audit fees (coefficient = 0.308), further substantiating the relevance of both signaling theory and information asymmetry theory (Scott, 2015; Spence, 1978). Manipulated accruals may serve as red flags to auditors, prompting greater scrutiny and higher audit costs. This result echoes the findings of Abbott et al. (2006) and Yu (2020), who also observed a risk premium associated with aggressive AEM. Notably, among the four accrual models employed, only the Kothari specification produced a statistically significant interaction with REM. By incorporating return on assets as a control variable, the Kothari model adjusts for performance-related distortions, yielding more accurate estimates of abnormal accruals (Kothari et al., 2005). Across all model specifications, the interaction term's t-statistic exceeded the 1.96 threshold only under the Kothari model. In contrast, the Jones, Modified Jones, and Kasznik models produced either weaker or insignificant results. While this does not imply the universal superiority of the Kothari model, the findings underscore its relatively greater explanatory power and sensitivity in detecting the interaction effects inherent in dual earnings management contexts.

From a practical perspective, the results are particularly relevant for auditors, regulators, and corporate governance bodies. Auditors are encouraged to develop more nuanced risk assessment procedures that account for the interplay between REM and AEM, especially in environments with limited transparency and high agency costs. Regulators may also consider developing guidelines that enhance disclosure related to both real and AEM activities, thereby equipping stakeholders with better tools to evaluate reporting integrity. Corporate boards and audit committees should also be vigilant in monitoring managerial decisions that might elevate informational risk and affect the reliability of financial statements.

These findings have several theoretical implications. First, they advance the current understanding of audit pricing by integrating the dual dimensions of earnings management, accrual-based and real, and highlighting their interactive effects. Unlike prior studies that analyze these constructs in isolation, this research offers a more comprehensive understanding of how simultaneous engagement in both types of earnings management influences auditors' pricing strategies. By demonstrating that the moderating effect of AEM is model-sensitive and only significant under the Kothari specification, this study contributes to ongoing methodological discussions in the literature on earnings management and audit risk.

Although Hypothesis 2 was formulated in a non-directional form to accommodate theoretical ambiguity in the literature, the empirical evidence demonstrates that the moderating effect of AEM is negative, indicating that higher levels of accrual-based earnings management attenuate the positive relationship between REM and audit fees. This finding clarifies the direction of the moderating mechanism proposed in the theoretical framework and aligns with the behavioral perspective, suggesting that auditors' attention may shift toward accrual anomalies when both forms of earnings management coexist.

6. Conclusion

This study examined the moderating effect of AEM on the relationship between REM and audit fees. First, consistent with Hypothesis 1 and the results reported in Table 4, REM exerts a positive and statistically significant effect on audit fees ($\beta = 0.580$, $t = 3.24$), confirming that auditors independently price REM as a distinct source of audit risk. Second, in line with H2, the negative coefficient of the interaction term (–2.640) indicates that AEM weakens this positive relationship. Together, these results highlight auditors' nuanced assessment of dual earnings management, while REM heightens perceived audit risk, concurrent AEM activity partially offsets the incremental pricing effect by complicating auditors' evaluation of overall reporting risk. Using a comprehensive dataset from firms listed on the Tehran Stock Exchange and employing multiple accrual models, the results revealed that REM significantly increases audit fees, indicating that auditors are sensitive to real manipulation practices. The concurrent application of real and AEM, referred to as dual earnings management, creates multifaceted risks that complicate auditors' efforts to evaluate the credibility and transparency of financial reports. Furthermore, AEM was found to moderate this relationship, but only under the Kothari specification, highlighting the model's sensitivity to this effect. These findings extend the audit pricing literature by providing a dual-dimensional perspective on earnings management and its implications for audit risk assessment. The study also emphasizes the importance of considering interaction effects between different forms of earnings management when evaluating the determinants of audit fees. Like all empirical research, this study has limitations. First, it excludes firms without audit fee disclosures and those in specific industries, which may limit the generalizability of the results. Second, the reliance on archival data restricts the ability to capture managerial intent behind earnings management. Future research could explore alternative proxies for earnings manipulation, employ interviews or surveys to gain deeper insights, or examine other emerging markets to validate the findings across institutional contexts.

Reference

1. Abbaszadeh, M. R., Ghanad, M. and Gholami Moghadam, F. (2017). Investigating the mediating effect of earnings management on the relationship between research and development expenses and audit service fees. *Accounting knowledge,* 8(2), pp. 85–104 (In Persian). https://doi.org/10.22103/jak.2017.1676
2. Abbott, L. J., Parker, S. and Peters, G. F. (2006). Earnings management, litigation risk, and asymmetric audit fee responses. *Auditing: A Journal of Practice & Theory*, 25(1), pp. 85-98. https://doi.org/10.2308/aud.2006.25.1.85
3. Aflatooni, A. (2018). Econometrics in financial and accounting research with EViews software. Termeh, Tehran, Iran (In Persian).
4. Alavi Tabari, S. H. and Haji Moradkhani, H. (2015). The relationship between audit quality and stock liquidity. *Financial Accounting and Auditing Research*, 27(7), pp. 92–110 (In Persian).

https://sanad.iau.ir/Journal/faar/Article/1073798

5. Alhadab, M. (2018). Abnormal audit fees and accrual and real earnings management: evidence from UK. *Journal of Financial Reporting and Accounting*, 16(3), pp. 395-416. https://doi.org/10.1108/JFRA-07-2017-0050
6. 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. https://www.doi.org/10.6018/rcsar.365091
7. Ashbaugh, H., LaFond, R. and Mayhew, B. W. (2003). Do nonaudit services compromise auditor independence? Further evidence. *The Accounting Review*, 78(3), pp. 611-639. https://doi.org/10.2308/accr.2003.78.3.611
8. Azad, A., Lari Dashtbayaz, M., Hadian, S. S. and Bagheri, M. (2023). The relationship between residual audit fees and real operating cash flow and real production expenditure smoothing. *Professional Auditing Research,* 4(13), pp. 156-174. https://doi.org/10.22034/jpar.2023.2008629.1196
9. Banimahd, B., Royaee, R. and Moshtagh Kahnamouei, M. (2022). Economic uncertainty and earnings management: evidence from listed companies at Tehran Stock Exchange. *Accounting and Auditing Research*, 14(54), pp. 5–20 (In Persian). https://doi.org/10.22034/iaar.2022.156641
10. Choi, A., Lee, E. Y., Park, S. and Sohn, B. C. (2022). The differential effect of accrual-based and real earnings management on audit fees: international evidence. *Accounting and Business Research*, 52(3), pp. 254-290. https://doi.org/10.1080/00014788.2021.1911779
11. Choi, A., Sohn, B. C. and Yuen, D. (2018). Do auditors care about real earnings management in their audit fee decisions?. *Asia-Pacific Journal of Accounting & Economics*, 25(1-2), pp. 21-41. https://doi.org/10.1080/16081625.2016.1231580
12. 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
13. Cohen, D. A. and Zarowin, P. (2010). Accrual-based and real earnings management activities around seasoned equity offerings. *Journal of Accounting and Economics*, 50(1), pp. 2-19. https://doi.org/10.1016/j.jacceco.2010.01.002
14. Darogheh Hazrati, F. and Pahlevan, Z. (2013). The Impact of the level of earnings management and audit fees on evidence from ISE. *Management Accounting*, 15(5), pp. 55–67 (In Persian). https://sanad.iau.ir/Journal/jma/Article/816549
15. Dechow, P. M., Sloan, R. G. and Sweeney, A. P. (1995). Detecting earnings management. *Accounting Review*, 70(2), pp. 193-225.
16. Egbunike, C. F., Igbinovia, I. M., Okafor, K. J. and Mmadubuobi, L. C. (2023). Residual audit fee and real income smoothing: evidence from quoted non-financial firms in Nigeria. *Asian Journal of Accounting Research*, 8(1), pp. 66-79. https://doi.org/10.1108/AJAR-01-2022-0006
17. Emadoddini, A. and Saeedi, P. (2019). Investigating the relationship between unusual audit fees and accruals management and actual earnings management. *Journal of Accounting and Management Vision*, 2(9), pp. 78-97 (In Persian). http://www.jamv.ir/article_90522.html
18. Frankel, R. M., Johnson, M. F. and Nelson, K. K. (2002). The relation between auditors' fees for nonaudit services and earnings management. *The Accounting Review,* 77(s-1), pp. 71-105. https://doi.org/10.2308/accr.2002.77.s-1.71
19. Greiner, A., Kohlbeck, M. J. and Smith, T. J. (2017). The relationship between aggressive real earnings management and current and future audit fees. *Auditing: A Journal of Practice & Theory*, 36(1), pp. 85-107. https://doi.org/10.2308/ajpt-51516
20. Haiderpoor, F. and Jafari, Y. (2016). The impact of product market structure on audit costs.

Financial Accounting and Auditing Research, 30(8), pp. 103–118 (In Persian). https://sanad.iau.ir/Journal/faar/Article/1073822
21. Hajiha, Z. and Chenari Bouket, H. (2016). Relationship between Earnings Management & Earnings Forecast Error: an Emphasis on the End & Interim Financial Statements. *Journal of Asset Management and Financing,* 4(2), pp. 29-48 (In Persian). https://doi.org/10.22108/amf.2016.20638
22. Healy, P. M. and Wahlen, J. M. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting Horizons*, 13(4), pp. 365-383. https://doi.org/10.2308/acch.1999.13.4.365
23. Hussien, L., Zraqat, O., Zureigat, Q., Al-Rawashdeh, H. A., Alrawashedh, N., Almutairi, A. and Alrashidi, M. (2024). The impact of audit quality on reducing earnings management practices in the Jordanian industrial companies. *In Harnessing AI, Machine Learning, and IoT for Intelligent Business: Volume 2 (pp. 569-578).* Cham: Springer Nature Switzerland, Cham, Switzerland.
24. Jones, J. J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research*, 29(2), pp. 193-228. https://doi.org/10.2307/2491047
25. Kasznik, R. (1999). On the association between voluntary disclosure and earnings management. *Journal of Accounting Research*, 37(1), pp. 57-81. https://doi.org/10.2307/2491396
26. Khoshkar, F., Rostami, S. and Khandeh, M. (2019). The relationship between audit fees, earnings quality, and financial reporting quality. *Journal of New Research Approaches in Management and Accounting*, 3(11), pp. 35–51 (In Persian). https://majournal.ir/index.php/ma/article/view/292
27. Kim, Y. and Park, M. S. (2014). Real activities manipulation and auditors' client-retention decisions. *The Accounting Review*, 89(1), pp. 367-401. https://doi.org/10.2308/accr-50586
28. Kordsholi, S. O. and Pajoohi, M. R. (2024). The Impact of the Risk Committee on the Quality Financial Reporting and Auditing Fees. *Financial Accounting and Auditing Research*, 61(16), pp. 127–150 (In Persian). https://doi.org/10.30495/faar.2024.709446
29. Kothari, S. P., Leone, A. J. and Wasley, C. E. (2005). Performance matched discretionary accrual measures. *Journal of Accounting and Economics*, 39(1), pp. 163-197. https://doi.org/10.1016/j.jacceco.2004.11.002
30. Krishnan, G. V., Sun, L., Wang, Q. and Yang, R. (2013). Client risk management: A pecking order analysis of auditor response to upward earnings management risk. *Auditing: A Journal of Practice & Theory,* 32(2), pp. 147-169. https://doi.org/10.2308/ajpt-50372
31. Kyriakou, M. I. (2025). The role of earnings management and audit fees in non-financial companies during crises. *Central Bank Review*, 25(2), A. 100200. https://doi.org/10.1016/j.cbrev.2025.100200
32. Meckling, W. H. and Jensen, M. C. (1976). Theory of the Firm. *Managerial behavior, agency costs and ownership structure*, *3*(4), pp. 305-360. https://doi.org/10.1007/978-94-009-9257-3_8
33. Mohammadrezaei, F. and Faraji, O. (2019). The dilemma of audit quality measuring in archival studies: critiques and suggestions for Iran's research setting. *Journal of Accounting and Auditing Review*, 26(1), pp. 87-122. https://doi.org/10.22059/acctgrev.2018.256781.1007884
34. Moradi, M., Safarpour, M. and Monfared, S. (2019). The relationship between earnings quality and audit fees with the dividend policy of listed companies on the Tehran Stock Exchange. *Journal of Accounting and Management Vision*, 2(12), pp. 1-16 (In Persian). http://www.jamv.ir/article_92554.html
35. Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of Accounting and Economics*, 42(3), pp. 335-370. https://doi.org/10.1016/j.jacceco.2006.01.002
36. Schipper, K. (1989). Commentary on earnings management. *Accounting Horizons*, 3(4), pp. 91-

102.
37. Schroeder, M. (2002). Senate panel seeks sweeping change for auditors. Wall Street Journal, New York, United States
38. Scott, W. R. (2015). Financial accounting theory. Pearson, United States: Canada Cataloguing, Canada.
39. 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
40. Sitanggang, R. P., Karbhari, Y., Matemilola, B. T. and Ariff, M. (2020). Audit quality and real earnings management: evidence from the UK manufacturing sector. *International Journal of Managerial Finance*, 16(2), pp. 165-181. https://doi.org/10.1108/IJMF-03-2018-0095
41. Sohn, B. C. (2011). Do auditors care about real earnings management in their audit fee decisions?, University of Macau, Macau. Available at SSRN 1899189. https://dx.doi.org/10.2139/ssrn.1899189
42. Spence, M. (1978). Job market signaling. In Uncertainty in economics (pp. 281-306). Academic Press, Cambridge, Massachusetts. https://doi.org/10.1016/B978-0-12-214850-7.50025-5
43. Sun, J., Lan, G. and Liu, G. (2014). Independent audit committee characteristics and real earnings management. *Managerial Auditing Journal*, 29(2), pp. 153-172. https://doi.org/10.1108/MAJ-05-2013-0865
44. Yu, C. (2020). Earnings management, audit costs and firm size. *In the 2020 International Conference on E-Commerce and Internet Technology (ECIT),* (pp. 90-96). IEEE, New York.
45. Zang, A. Y. (2012). Evidence on the trade-off between real activities manipulation and accrual-based earnings management. *The Accounting Review*, 87(2), pp. 675-703. https://doi.org/10.2308/accr-10196
46. Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research*, 22(1), pp. 59-82. https://doi.org/10.2307/2490859