JOURNAL OF MANAGEMENT AND BUSINESS SOLUTIONS

Simulation of Rational Investment Decision-Making Based on Earnings Quality and Its Impact on Future Stock Returns in the Tehran Stock Exchange

1. Morteza. Hadadi : Ph.D. student, Department of Accounting, Bo.c., Islamic Azad University, Borujerd, Iran
2. Alireza. Ghiyasvand *: Assistant Professor, Department of Accounting, Bo.c., Islamic Azad University, Borujerd, Iran
3. Farid. Sefaty: Assistant Professor, Department of Accounting, Bo.c., Islamic Azad University, Borujerd, Iran

*corresponding author's email: 4130700286@iau.ac.ir

ABSTRACT

This study was conducted with the aim of simulating investors' rational decision-making processes based on earnings quality and examining its impact on future stock returns in the Tehran Stock Exchange. The research also investigated the moderating role of behavioral biases and proposed an integrated framework based on game theory and Bayesian inference for portfolio optimization. Data from 129 listed firms during the period 2001–2023 (2,193 firm-year observations) were selected using a systematic deletion method. The integrated methodology included dynamic generalized method of moments (GMM) regression to control for endogeneity; artificial neural networks, LSTM, and Transformer models for simulating decision-making; SHAP analysis for interpretability; and Monte Carlo simulation with 100,000 iterations for portfolio optimization. The dimensions of earnings quality (accrual quality: 0.315, earnings persistence: 0.271, predictability: 0.203) had a significant and positive effect on future returns. Adding behavioral variables increased R^2 from 47.36% to 52.47% and reduced RMSE by 13.44%. The Friedman test ($\chi^2 = 187.45$, $p < 0.0001$) confirmed the significant moderating role of behavioral factors. The Multi-head Transformer model demonstrated the best predictive performance ($R^2 = 85.12\%$). The optimized portfolio produced a return of 26.12% with a Sharpe ratio of 1.351, which was 83.6% better than the market index. Therefore, the results indicate that integrating fundamental factors (earnings quality) and behavioral factors within a unified game-theoretic framework significantly enhances the predictive power of future returns. Investors, in addition to traditional financial analysis, should pay particular attention to reporting-quality signals and behavioral indicators of the market.

Keywords: earnings quality, future returns, rational decision-making simulation, behavioral biases, game theory, deep learning

Introduction

The question of how investors process accounting information and transform it into buy–hold–sell decisions has become increasingly complex in the era of high-frequency trading, alternative data, and algorithmic portfolio construction. Although classical finance assumes rational agents who update beliefs based on earnings and cash-flow news, a growing body of evidence suggests that the market's reaction to accounting signals is shaped jointly by earnings quality, behavioral biases, and the modeling capacity of modern prediction tools (1, 2). In emerging markets such as Iran, where institutional frictions, financing constraints, and information asymmetries are more pronounced, the quality of financial reporting may play an even more pivotal role in the pricing of risk and the

Article history:
Received 04 April 2025
Revised 30 May 2025
Accepted 02 June 2025
Published online 10 June 2025

How to cite this article:
Hadadi, M., Ghiyasvand, A., & Sefaty, F. (2025). Simulation of Rational Investment Decision-Making Based on Earnings Quality and Its Impact on Future Stock Returns in the Tehran Stock Exchange. *Journal of Management and Business Solutions, 3*(3), 1-26. https://doi.org/10.61838/jmbs.85

formation of future returns (3, 4). Understanding how investors *should* behave under rational expectations, and how they *actually* behave in the presence of sentiment and noise, requires an integrated framework that goes beyond traditional linear models and incorporates machine learning and behavioral finance perspectives (5, 6).

Earnings quality has long been recognized as a fundamental determinant of the decision-usefulness of financial statements and of the efficiency of capital allocation in securities markets (7). Prior studies document that higher-quality earnings—characterized by persistent, predictable, and accruals-based measures that faithfully represent underlying performance—are associated with more accurate valuation, lower cost of capital, and more efficient investment decisions (1, 8). In the context of emerging markets, deficiencies in earnings quality can exacerbate agency problems and information asymmetry, leading investors to misprice risk and misallocate capital across projects and sectors (9). Empirical evidence indicates that when earnings are subject to discretionary accruals, real activity manipulation, or weak conservatism, the market's ability to infer sustainable cash flows and future returns is impaired, thereby undermining the predictive content of accounting numbers (10, 11). Consequently, earnings quality is not merely an accounting construct; it is a strategic variable that shapes investors' beliefs about firm behavior and future performance, particularly in environments with limited alternative information sources (7, 8).

Recent research in the Iranian capital market underscores this challenge by showing that financial reporting quality and governance-sensitive accounting information affect firms' investment policies and the credibility of their forward-looking disclosures (8, 12). High-quality management forecasts and transparent earnings streams appear to reduce under- and over-investment, enhance investment efficiency, and strengthen the role of accounting information as a governance mechanism in monitoring managerial decisions (12, 13). At the same time, weak reporting environments and opportunistic earnings management—whether accrual-based or real—distort capital structure choices and undermine the disciplining role of capital markets, particularly when combined with low-quality disclosure and inadequate investor protection (11, 14). These findings highlight the need to embed earnings quality explicitly in models of investor decision-making and portfolio construction, especially when markets face structural financing bottlenecks and sector-specific vulnerabilities such as those observed in tourism and production financing in Iran (3, 4).

Parallel to this line of inquiry, a rich literature on return forecasting has evolved, ranging from classical time-series approaches to sophisticated machine learning architectures. Traditional models such as Box–Jenkins, exponential smoothing, and Bayesian methods have been widely applied to forecast stock returns in the Tehran Stock Exchange, with mixed evidence regarding their out-of-sample accuracy and stability across market regimes (15, 16). More recent studies incorporate comprehensive sets of financial ratios and risk indicators to predict negative price shocks and downside risk, emphasizing the predictive content of leverage, profitability, liquidity, and market-based variables (9, 17). In addition, knowledge-domain analysis and Delphi–fuzzy techniques have been used to map the conceptual space of stock return predictors and prioritize key drivers that should enter forecasting models in emerging markets (9). However, the linearity assumptions underlying many of these approaches may be too restrictive in the presence of non-linear interactions among earnings quality, growth opportunities, and risk proxies, especially under volatile macro-financial conditions.

Machine learning provides a powerful alternative to these traditional tools by allowing flexible functional forms, automatic feature selection, and the exploitation of high-dimensional predictor spaces. Empirical asset-pricing research demonstrates that tree-based methods, neural networks, and other learning algorithms can extract non-linear patterns in the cross-section of returns that standard linear models fail to capture (2). In international markets, Long Short-Term Memory (LSTM) networks and multilayer perceptrons have been applied to forecast stock return movements using both price-based technical indicators and recent fundamental information, often yielding higher predictive accuracy than benchmark models (18, 19). Research on the Indonesian and S&P 500 markets shows that incorporating up-to-date accounting variables and macro signals into deep architectures improves both directional accuracy and risk-adjusted performance, particularly when models are carefully tuned and validated (18, 19). Studies focused on P/E ratios and panel data settings further integrate LSTM with sentiment indicators, underscoring the importance of behavioral dimensions in modern return forecasting systems (20).

The Iranian context has also begun to benefit from such techniques. Studies comparing ensemble classifiers, nearest-neighbor methods, decision trees, and other learning machines document meaningful improvements in predicting returns from stock price changes relative to traditional benchmarks, albeit with varying robustness across sectors and time horizons (21, 22). Other works explore the use of distributional assumptions, such as Laplace-based modeling of aggregate market returns and volatility, to derive more accurate probabilistic forecasts for return distributions (23). Importantly, evidence suggests that combining machine learning approaches with domain knowledge from accounting and finance—rather than treating them as black boxes—enhances both interpretability and acceptance by practitioners and regulators (22, 24). Nevertheless, the majority of these models still treat earnings quality and behavioral factors as peripheral inputs rather than as core structural elements of investor decision-making.

At the same time, behavioral finance has highlighted that investors do not always behave as fully rational Bayesian updaters. Meta-analytic evidence reveals systematic gaps between ex ante predictions and realized performance, driven by information overload, biased processing of signals, and overreliance on salient but noisy cues, particularly in the evaluation of new ventures and high-uncertainty projects (6). In the Tehran Stock Exchange, empirical research demonstrates that investor sentiment—proxied by trading activity, market-wide indicators, and survey-based measures—affects returns, cash flows, discount rates, and overall firm performance beyond what is explained by fundamentals alone (25). Sentiment-driven mispricing can be particularly pronounced in markets where short-selling constraints, limited arbitrage, and concentrated ownership structures weaken the corrective forces of rational traders (3, 25). International evidence on sentiment-augmented return prediction, such as the use of sentiment to model P/E ratios, similarly suggests that emotional and psychological factors shape valuation ratios and market dynamics in ways that must be explicitly modeled rather than treated as exogenous noise (20).

Earnings management is a key mechanism through which managerial behavior interacts with both fundamental and behavioral dimensions of the market. Research distinguishes between accrual-based and real earnings management, showing that both forms can distort firms' apparent profitability, risk profile, and capital structure decisions (11, 14). In periods of financial stress, managers may resort more intensively to income-smoothing and opportunistic accounting techniques, with significant implications for investors' ability to assess sustainability of earnings and predict future returns (14, 17). Evidence on real earnings quality and its interaction with dividend policy further indicates that payout decisions can moderate or amplify the market's perception of earnings credibility and future firm value (10). From a capital-market perspective, the challenge is not only to detect low-quality reporting ex post, but to embed expectations about earnings quality into ex ante decision models that determine portfolio weights and risk exposures (1, 8).

This need for integrated decision frameworks has spurred interest in combining financial reporting quality, investment efficiency, and decision-support systems into holistic models for investors and policymakers. Studies in the Iranian setting have designed conceptual and empirical models that link financial reporting quality to investor decision-making quality and investment efficiency, often using mixed-methods approaches to capture both qualitative expert judgment and quantitative performance metrics (13, 26). Other research has developed optimal decision-making frameworks that explicitly integrate artificial intelligence, transparency of financial reporting, and multi-criteria evaluation to help investors process complex accounting and market information under uncertainty (5). Network data envelopment analysis (DEA) models have been proposed to assess the informational efficiency of reporting units, emphasizing that efficiency evaluation itself depends critically on the reliability and structure of the underlying accounting data (27). These contributions suggest that investor-support systems must be built on a solid accounting-information foundation while leveraging advanced analytics to prioritize relevant signals and mitigate biases.

However, the rapid diffusion of artificial intelligence into accounting, auditing, and finance raises new ethical and governance challenges. On the one hand, AI-based systems can enhance the speed, accuracy, and consistency of financial analyses, anomaly detection, and forecasting; on the other hand, they may embed opaque decision rules, reproduce historical biases, or erode accountability if not properly designed and regulated (24). Studies on AI-driven decision-making stress the importance of transparency, explainability, and ethical safeguards, especially when algorithms are used to support high-stakes investment and risk-management decisions (5, 24). In emerging capital markets facing structural financing gaps and policy constraints, such as production and tourism finance in Iran, the potential benefits of AI-based forecasting and optimization must be weighed against concerns about data quality, model risk, and institutional readiness (3, 4). These debates reinforce the need for simulation-based approaches that not only maximize predictive accuracy, but also remain interpretable and aligned with rational economic principles.

Against this backdrop, recent work has begun to explicitly model rational decision-making using game-theoretic and Bayesian frameworks combined with simulation and machine learning. For instance, illustrative simulations have been developed in which earnings quality drives rational portfolio rebalancing, and investors update their beliefs about future returns conditional on observable accounting signals (10, 28). In parallel, empirical studies on forecasting stock returns with financial and regulatory criteria using machine learning in the Iranian market highlight the need to jointly consider institutional constraints, regulatory signals, and firm-level accounting information (22). Research on the feasibility or impossibility of predicting stock prices in specific industries—such as petro-refining—further emphasizes that predictability may be sector-dependent and shaped by the interaction of fundamentals, market microstructure, and policy interventions (29). These developments suggest that building robust decision-support systems for investors requires integration of earnings quality metrics, behavioral proxies, institutional features, and advanced learning architectures within a unified simulation-based framework.

Despite these advances, several important gaps remain. First, there is limited evidence on how investors in the Tehran Stock Exchange would behave under a *rational* benchmark that explicitly conditions on multidimensional earnings quality measures while also acknowledging the presence of sentiment and trading-behavior signals (20, 25). Second, the literature has not fully explored how different families of machine learning models—such as multilayer perceptrons, LSTM networks, and Transformer architectures—compare in forecasting future returns when they are fed with both fundamental and behavioral variables derived from the same market (2, 18, 19). Third, the interaction between earnings quality, investment efficiency, and macro-level financing constraints in Iran's capital market suggests a need for portfolio-optimization models that explicitly encode rational decision rules based on quality-adjusted earnings and sentiment-adjusted risk assessments (3, 4, 13). Finally, while prior studies have assessed real earnings management, capital structure, and governance considerations in isolation, there is still insufficient understanding of how these mechanisms jointly affect the mapping from accounting signals to expected returns within a rational yet behaviorally enriched decision framework (11, 14, 17).

In response to these gaps, the present study focuses on the Tehran Stock Exchange and develops an integrated simulation framework that combines dynamic panel GMM, artificial neural networks, LSTM and Transformer models, game-theoretic and Bayesian reasoning, and interpretable AI tools to model rational investor decision-making based on earnings quality while explicitly incorporating behavioral biases and portfolio optimization criteria (1, 2, 5, 21, 28). Accordingly, the aim of this study is to simulate rational investor decision-making based on multidimensional earnings quality measures and behavioral signals and to examine their joint impact on future stock returns and optimal portfolio allocation in the Tehran Stock Exchange.

Methods and Materials

This study employs a mixed-method methodological design. To test the hypotheses and analyze the role of earnings quality in future returns while considering behavioral biases, a combination of dynamic panel econometrics using two-stage least squares (2SLS) and an artificial neural network (ANN)–based simulation of rational decision-making is applied. The statistical population consists of all firms listed on the Tehran Stock Exchange. The time period spans continuously from 2001 to the end of 2023 to ensure sufficient data for dynamic panel modeling and to capture different economic cycles in the market. Ultimately, the sample was selected using a systematic elimination method, and firms that met four main criteria (continuous activity, non-financial and non-holding firms, stable fiscal year ending in March, and data availability) were included as the final sample. Accordingly, the sample-selection process is presented in the following:

Total number of listed firms at the end of 2023: 845

Firms not active during the 2001–2023 period: (256)

Firms listed after 2001: (340)

Firms classified as holdings, investment, financial intermediaries, banks, or leasing companies: (64)

Firms that changed fiscal year or did not end the fiscal year in March: (54)

Firms with unavailable data during the study period: (2)

Final sample firms: 129

After applying all the above criteria, 129 firms remained as the screened population, all of which were selected as the study sample.

Following the presentation of the population, sample, and study period, the analytical models and techniques are specified. Based on the seven research hypotheses and the multidimensional nature of the variables, hypotheses 1–4 (related to the direct effects of fundamental financial factors) are tested using generalized method of moments (GMM) regression to control endogeneity and heteroskedasticity. Furthermore, to test hypotheses 5–7, which focus on behavioral biases and rational decision-making, advanced data-mining approaches—such as ANN-based simulation and stepwise regression—are used to determine not only the effects but also the priority and relative weights of financial and behavioral variables in the stock-pricing process. In alignment with the model of Silaljaja and Jasman (2024), the models are specified as follows:

Regression Model (1): Testing Hypotheses 1–4 (Financial Factors)

This model estimates the direct effect of fundamental (financial) factors on future stock returns.

$$\begin{aligned} R_{i,t+1} \text{ and } E_{i,t+1} \quad &= \beta_0 + \beta_1 R_{i,t} \text{ and } ROE_{i,t} + \beta_2 AQ_{i,t} + \beta_3 EP_{i,t} + \beta_4 EPr_{i,t} \\ &+\beta_5 ESMO_{i,t} + \beta_6 ER_{i,t} + \beta_7 ECON_{i,t} + \beta_8 EPS_{i,t} + \beta_9 DPS_{i,t} \\ &+\beta_{10} DY_{i,t} + \beta_{11} DK_{i,t} + \beta_{12} TAXAVO_{i,t} + \beta_{13} ETR_{i,t} \\ &+\beta_{14} AGR_{i,t} + \beta_{15} SGR_{i,t} + \beta_{16} GO_{i,t} + \beta_{17} SIZE_{i,t} \\ &+\beta_{18} ROA_{i,t} + \beta_{19} ROI_{i,t} + \beta_{20} FIRMAGE_{i,t} + \eta_i + \delta_t + \varepsilon_{i,t} \end{aligned} \tag{1}$$

Regression Model (2): Testing Hypothesis 5 (Behavioral Factors)

In this model, four behavioral variables are added to Model (1).

$$\begin{aligned} R_{i,t+1} \text{ and } E_{i,t+1} \quad &= \beta_0 + \beta_1 R_{i,t} \text{ and } ROE_{i,t} + \beta_2 AQ_{i,t} + \beta_3 EP_{i,t} + \beta_4 EPr_{i,t} \\ &+\beta_5 ESMO_{i,t} + \beta_6 ER_{i,t} + \beta_7 ECON_{i,t} + \beta_8 EPS_{i,t} + \beta_9 DPS_{i,t} \\ &+\beta_{10} DY_{i,t} + \beta_{11} DK_{i,t} + \beta_{12} TAXAVO_{i,t} + \beta_{13} ETR_{i,t} \\ &+\beta_{14} AGR_{i,t} + \beta_{15} SGR_{i,t} + \beta_{16} GO_{i,t} + \beta_{17} SIZE_{i,t} \\ &+\beta_{18} ROA_{i,t} + \beta_{19} ROI_{i,t} + \beta_{20} FIRMAGE_{i,t} \\ &+\gamma_1 RSI_{i,t} + \gamma_2 P\text{-}Line_{i,t} + \gamma_3 Sentiment_{i,t} + \gamma_4 TradingBehavior_{i,t} + \eta_i + \delta_t + \varepsilon_{i,t} \end{aligned} \tag{2}$$

The purpose of Model (2) is to compare the predictive accuracy (RMSE) of Model (1) and Model (2) using the Diebold–Mariano test.

Regression Model (3): Testing the Interactive (Moderating) Effect

This model tests the moderating effect of behavioral factors (B) on the relationship between financial factors (X) and future returns (R). It tests hypotheses 5 and 6 using regression (not ANN), and complements previous models:

$$\begin{aligned} R_{i,t+1} \text{ and } E_{i,t+1} \quad &= \beta_0 + \beta_1 R_{i,t} \text{ and } ROE_{i,t} + \sum_{j=1}^{15} \beta_j (Financial\_Variables)_{i,t} \\ &+ \sum_{k=1}^{4} (Behavioral\_Variables)_{i,t} + \sum_{l=1}^{P} \lambda_l (X \times B)_{i,t} \\ &+ \sum_{m=1}^{4} \delta_m (Control\_Variables)_{i,t} + \eta_i + \delta_t + \varepsilon_{i,t} \end{aligned} \tag{3}$$

The purpose of Model (3) is to examine the statistical significance of the interaction coefficients (λ_l). If these coefficients are significant, it indicates that behavioral factors moderate the relationship between financial factors and future returns.

Therefore, to test hypothesis 6 and evaluate the influence of investors' behavioral biases on their rational decision-making, an ANN-based simulation and a regression decision-tree (RDT) algorithm are used. This process is performed in several stages to derive the relative weights and importance of variables:

Stage 1: Determining the Importance of Financial Variables

All financial variables (X)—including earnings quality, dividend policy, performance, growth, and other fundamental variables—are used as ANN inputs. Behavioral biases are not included at this stage.

The ANN is trained with financial variables, and the importance coefficient of each variable in rational decision-making is obtained.

To evaluate importance, sensitivity analysis and integration with the regression decision-tree algorithm are used.

Model structure at this stage:

$$Q_{i,t} = f(X_{i,t})$$

where **Q** represents the rational decision-making index of investor *i* in year *t*, and **X** represents each financial variable.

Stage 2: Including Behavioral Variables

In this stage, behavioral variables (B)—including Relative Strength Index (RSI), psychological line index, investor sentiment, and trading behavior—are added as behavioral independent variables.

The simulated model becomes:

$$Q_{i,t} = f(X_{i,t}, B_{i,t})$$

The ANN is retrained using the expanded dataset, and the relative importance of all financial and behavioral variables is recalculated.

Stage 3: Statistical Testing of Hypothesis 6

To compare variable importance before and after adding behavioral variables, the rank of each variable is recorded and changes are examined.

The nonparametric Friedman test is used to test the statistical significance of these changes.

If the Friedman test shows that financial-variable rankings change significantly in the presence of behavioral variables, hypothesis 6 is supported.

In this study, in addition to regression models and machine learning algorithms, a theoretical framework based on game theory is used to simulate the rational decision-making process of investors. This approach models the behavior of the firm and the investor as a strategic game in which the investor attempts to predict the firm's next move (the true quality of reporting) by analyzing financial signals and thereby maximize his or her payoff. Therefore, in order to gain a deeper understanding of why rational decision-making is based on signals sent by the firm, a game-theoretic framework is employed. This approach models the interaction between the firm (reporter) and the rational investor (decision-maker) as a strategic game, the scenario of which is presented in Table 1 as follows.

Table 1. Strategic Game Scenario

Player	Strategies (Signals)
Firm (Player 1)	1. Quality of manipulation activity (accounting compliance): High Compliance or Low Compliance/Manipulation 2. Tax management: Proper management (high compliance) or aggressive tax management (low compliance)
Investor (Player 2)	Investment decision: Buy (high confidence) or Sell (low confidence/high risk)

The investor's decision-making process begins with a perceptual mapping in which the investor evaluates two key signals: business growth (as a proxy for the level of compliance with accounting standards) and the firm's tax compliance. The interaction of these two factors shapes the investor's perception of the firm's intent (a prudent versus an opportunistic approach).

A rational investor considers high tax compliance, regardless of the firm's growth level, as a positive signal indicating a prudent managerial approach.

In contrast, low tax compliance (aggressive tax avoidance) is perceived as an opportunistic incentive and leads to a negative perception of reporting quality.

To formalize this strategic interaction, a game matrix is used in which the firm's main strategies are the quality of manipulation activity and tax management. The game matrix for investment decision-making is shown in Table 2 as follows.

Table 2. Game Matrix for Investment Decision-Making

Manipulation activity	High compliance	Low compliance
Proper tax management (high compliance)	High confidence and favorable buy position (Quadrant I: win–win)	Unfavorable buy position (Quadrant III: low probability)
Aggressive tax management (low compliance)	Sell position (Quadrant II: low probability)	Unfavorable sell position (Quadrant IV: lose–lose and high probability)

Quadrant I (win–win): High accounting compliance combined with proper tax management is a strong signal of transparency and high quality and creates the most favorable buy position.

Quadrant IV (lose–lose): Low accounting compliance (earnings manipulation) combined with aggressive tax management represents the worst possible state, indicating high risk and lack of transparency, and results in a definite sell position.

Probabilistic Modeling for Buy and Sell Positions

To quantify decisions based on the game matrix, Bayes' rule is used to compute the probability of taking a "buy" or "sell" position given the observed signals.

a) Formula for the probability of a buy position:

This formula calculates the probability that a firm is favorable (high manipulation-activity quality and proper tax management) given the received signals:

$$P(DTAQ,HI \mid MAQ,H) = \frac{P(MAQ,H \mid DTAQ,HI) \times P(DTAQ,HI)}{P(MAQ,H \mid DTAQ,HI) \times P(DTAQ,HI) + P(MAQ,H \mid DTAQ,Lo) \times P(DTAQ,Lo)}$$

b) Formula for the probability of a sell position:

This formula calculates the probability that a firm is unfavorable (low manipulation-activity quality) even when an apparently positive signal (proper tax management) is observed:

$$P(DTAQ,HI \mid MAQ,L) = \frac{P(MAQ,L \mid DTAQ,HI) \times P(DTAQ,HI)}{P(MAQ,L \mid DTAQ,HI) \times P(DTAQ,HI) + P(MAQ,L \mid DTAQ,Lo) \times P(DTAQ,Lo)}$$

In the above relations, $P(MAQ,H)$is the prior probability that the firm has high manipulation-activity quality; $P(MAQ,L)$is the prior probability that the firm has low manipulation-activity quality; $P(DTAQ,HI)$is the probability of observing proper (high) tax management; and $P(DTAQ,Lo)$is the probability of observing aggressive (low) tax management.

Portfolio Optimization Modeling

Finally, the outputs of the above probabilistic model are used as inputs for a portfolio optimization model. The goal of this model is the optimal allocation of capital among the four largest firms in the market based on financial reporting quality criteria.

Objective function:

$$\text{Maximize } Z = D_1X_1 + D_2X_2 + D_3X_3 + D_4X_4$$

where X_1–X_4 are the percentages of capital allocated to each of the four top firms, and D_1–D_4 are decision coefficients (for example, the probability of a buy position calculated from Bayes' formula) for each firm. This optimization is carried out under the following qualitative constraints to ensure that the final portfolio consists of firms that meet at least minimum reporting standards:

Quality of manipulation activity:

$$\delta_1 X_1 + \delta_2 X_2 + \delta_3 X_3 + \delta_4 X_4 < \text{market average}$$

Quality of discretionary tax accruals:

$$\mu_1 X_1 + \mu_2 X_2 + \mu_3 X_3 + \mu_4 X_4 < \text{market average}$$

Dividends:

$$\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 X_4 > \text{market average}$$

Sales growth:

$$\zeta_1 X_1 + \zeta_2 X_2 + \zeta_3 X_3 + \zeta_4 X_4 > \text{market average}$$

The coefficients δ, μ, α, and ζindicate the sensitivity of each qualitative criterion for the respective firm. This model ensures that investment is made only in firms that simultaneously have manipulation and discretionary tax-accrual quality measures below the market average (indicating higher quality), and dividends and sales growth above the market average. This integrated framework enriches the research method by accurately simulating the behavior of a rational investor.

In relations (1), (2), and (3), the research variables and how each is measured are specified as shown in Table 3.

Table 3. Research Variables with Two Dependent Variables

Variable category	Symbol	Variable definition	Measurement and calculation
Main dependent variable	$R_{i,t+1}$	Future stock returns	Ratio of the change in closing price of stock *i* at the end of year *t+1* plus cash dividend distributed, divided by the closing price at the end of year *t*.
Dependent variable (robustness)	$E_{i,t+1}$	Market value based on shareholders' equity	Ratio of market value of shareholders' equity of firm *i* at the end of year *t+1* to the book value of shareholders' equity at the end of year *t*.
Earnings quality (EQ)	$AQ_{i,t}$	Accrual quality	Absolute value of discretionary accruals (DA) based on adjusted models, divided by total assets.
	$EP_{i,t}$	Earnings persistence	Regression coefficient $\rho$from the model $E_{i,t} = \rho_0 + \rho_1 E_{i,t-1} + \varepsilon_{i,t}$.
	$EPr_{i,t}$	Earnings predictability	Inverse of the standard deviation of residuals from the time-series earnings model over the previous five years.
	$ESMO_{i,t}$	Earnings smoothing	Ratio of the standard deviation of net income to the standard deviation of sales.
	$ER_{i,t}$	Earnings relevance	Coefficient of determination $R^2$from the regression of stock price on earnings per share ($P_{i,t} = \beta_0 + \beta_1 EPS_{i,t} + \varepsilon$).
	$ECON_{i,t}$	Earnings conservatism	Negative coefficient of the earnings-change variable in asymmetric models (such as the adjusted Basu model).
Profitability / financial	$EPS_{i,t}$	Earnings per share	Net income in year *t* divided by the number of ordinary shares.
	$DPS_{i,t}$	Dividend per share	Cash dividend distributed per share.
	$DY_{i,t}$	Dividend yield	Ratio of dividend per share (DPS) to the closing share price.

	$DK_{i,t}$	Type of dividend payment	Dummy variable: 1 if payment is in cash; 0 otherwise.
	$TAXAVO_{i,t}$	Tax avoidance	Ratio of the difference between statutory tax and cash tax paid to total assets.
	$ETR_{i,t}$	Effective tax rate	Ratio of tax expense to profit before tax.
	$AGR_{i,t}$	Asset growth	Ratio of the increase in total assets to total assets at the beginning of the year.
	$SGR_{i,t}$	Sales growth	Ratio of the increase in net sales to net sales in year $t{-}1$.
	$GO_{i,t}$	Growth opportunities	Ratio of market value of shareholders' equity to book value of shareholders' equity (market-to-book).
Behavioral factors (B)	$RSI_{i,t}$	Relative Strength Index	Value of the RSI indicator at the end of year t.
	$P\text{-}Line_{i,t}$	Psychological line index	Value of P-Line (ratio of the number of up days to total trading days).
	$Sentiment_{i,t}$	Investor sentiment	Ratio of annual turnover of the firm's stock i.
	$TradingBehavior_{i,t}$	Trading behavior	Logarithm of abnormal trading volume.
Control variables	$SIZE_{i,t}$	Firm size	Natural logarithm of total assets.
	$FIRMAGE_{i,t}$	Firm age	Natural logarithm of the number of years of the firm's activity since listing.
	$ROA_{i,t}$	Return on assets	Ratio of net income to total assets.
	$ROI_{i,t}$	Return on investment	Ratio of operating profit (EBIT) to total assets.
Interaction terms	$X \times B$	Interaction effects (λ)	Product of financial variables (X) and behavioral variables (B).

Findings and Results

Before implementing the econometric models and neural networks, the research dataset was subjected to a comprehensive preprocessing procedure to ensure the validity and accuracy of the results. In the first step, in order to preserve the maximum number of observations and avoid reducing the power of statistical tests, the median imputation method was used instead of deleting rows with missing data. In this method, all missing values in continuous and financial variables were replaced with the median of the same variable over the study period. This procedure effectively resolved the problem of incomplete observations without introducing serious bias into the mean of the distributions. In the next step, to neutralize the adverse effects of outliers arising from recording errors or very rare financial events on regression estimates, a winsorization (boundary adjustment) technique was employed. After this adjustment, all continuous variables entered the normalization stage. For this purpose, Z-score standardization was used, which scales the variables so that they have a mean of zero and a standard deviation of one. This standardization, which ensures homogeneity of measurement scales, is a critical condition for preventing variables with high variance from dominating the regression analysis and the inputs of neural networks. Finally, the fully cleaned, adjusted, and normalized dataset became ready for subsequent descriptive and inferential analyses. Descriptive statistics are reported in Table 4.

Table 4. Descriptive Statistics of Research Variables

Variable	Symbol	Mean	Median	Maximum	Minimum	Standard deviation	Skewness	Kurtosis	Number of observations
Future returns	R	-0.0137	-0.0441	0.9955	-0.9665	0.2064	1.2365	7.9611	2,193
Accrual quality	AQ	-0.0421	-0.0730	0.9922	-0.4973	0.1737	1.6795	8.0801	2,193
Earnings persistence	EP	-0.0488	-0.1145	0.6927	-0.5112	0.1996	1.5976	5.3440	2,193
Earnings predictability	EPR	-0.0286	-0.0578	0.9082	-0.7273	0.2105	1.0056	4.8557	2,193
Earnings conservatism	ECON	-0.0488	-0.1145	0.6927	-0.5112	0.1996	1.5976	5.3440	2,193
Earnings smoothing	ESMO	-0.0027	-0.0048	0.9965	-0.9937	0.0680	0.4968	117.8770	2,193

Earnings per share	EPS	-0.0187	-0.0508	0.9960	-0.9745	0.1413	2.8422	18.8266	2,193
Dividend per share	DPS	0.0102	0.0039	0.9251	-0.7929	0.2906	0.0516	2.6983	2,193
Dividend yield	DY	-0.0343	-0.0719	0.9935	-0.4079	0.1664	2.7386	14.0010	2,193
Type of dividend payment	DK	0.20	0.0000	1.0000	0.0000	0.4005	1.4950	3.2350	2,193
Tax avoidance	TAXAVO	-0.0081	-0.0095	0.7063	-0.0673	0.0288	13.8206	262.6250	2,193
Volatility of effective tax rate	ETRV	0.0073	0.0058	0.4961	-0.7058	0.0308	-3.6578	214.6100	2,193
Growth opportunities	GO	-0.0198	0.0144	0.9779	-0.8255	0.2798	0.6252	4.8091	2,193
Asset growth	AGR	-0.0235	-0.0555	0.9917	-0.6223	0.2090	1.0594	5.4288	2,193
Sales growth	SGR	0.0077	0.0035	0.7252	-0.6944	0.2755	-0.0113	2.4352	2,193
Relative Strength Index	RSI	-0.0414	-0.0921	0.9792	-0.4177	0.1880	2.3920	9.6083	2,193
Psychological line	PSY	0.0286	0.0405	0.2226	-0.8785	0.0872	-5.7748	42.9579	2,193
Investor emotions	ATR	0.08704	0.11183	78.1790	-23.8670	0.05004	0.41031	21.1120	2,193
Trade imbalance	BSI	-0.0172	-0.0125	0.9068	-0.9420	0.3340	-0.0664	2.8793	2,193
Corporate image	FRIMAGE	0.0114	0.0495	0.8594	-0.8977	0.3060	-0.3428	3.3415	2,193
Return on assets	ROA	-0.0270	-0.0910	0.8050	-0.6408	0.2781	0.6813	2.8477	2,193
Return on equity	ROE	0.0068	0.0067	0.7383	-0.9903	0.0908	-1.2434	29.7871	2,193
Firm size	SIZE	-0.0235	-0.0631	0.8920	-0.8517	0.2923	0.4346	3.0559	2,193
Return on investment	ROI	-0.0248	-0.0606	0.8081	-0.6511	0.1803	0.7702	4.1036	2,193

The descriptive statistics table shows that the mean future stock return is negative and equal to -0.0137, indicating the overall weak performance of the sample during the study period. Earnings quality variables such as accrual quality, earnings persistence, and earnings conservatism also have negative means, suggesting low financial reporting quality among the sample firms. The high skewness and kurtosis values in some variables, such as earnings smoothing with kurtosis of 117.877 and earnings per share with kurtosis of 18.8266, indicate the presence of outliers and non-normal data distributions, which justifies the use of robust econometric methods. Behavioral and market variables also exhibit interesting patterns; the type of dividend payment, with a mean of 0.20, shows that only 20% of the observations involve cash dividend payments. Market sentiment variables such as the Relative Strength Index and psychological line have negative means, reflecting a generally negative psychological atmosphere in the market. The relatively high standard deviation in variables such as corporate image (0.3060) and trade imbalance (0.3340) indicates substantial heterogeneity and dispersion in the behavior of investors and firms, creating an appropriate opportunity to analyze differential effects.

Table 5. Diagnostic Test Results for GMM Models

Diagnostic test	Model (1) – Financial factors	Model (2) – Financial + behavioral factors	Model (3) – With interaction effects
Instrument validity tests			
Sargan statistic	28.45 (p = 0.241)	31.23 (p = 0.198)	34.67 (p = 0.156)
Hansen statistic	26.78 (p = 0.289)	29.56 (p = 0.234)	32.41 (p = 0.187)
Autocorrelation tests			
AR(1)	-3.456*** (p = 0.001)	-3.612*** (p = 0.000)	-3.734*** (p = 0.000)

AR(2)	-1.234 (p = 0.217)	-1.187 (p = 0.235)	-1.156 (p = 0.248)
Heteroskedasticity tests			
White test	234.56*** (p = 0.000)	267.89*** (p = 0.000)	298.45*** (p = 0.000)
Breusch–Pagan test	187.34*** (p = 0.000)	213.67*** (p = 0.000)	241.23*** (p = 0.000)
Endogeneity test	45.67*** (p = 0.000)	52.34*** (p = 0.000)	58.91*** (p = 0.000)
Number of instruments	42	48	56
F-statistic	127.45***	143.67***	156.89***

The diagnostic tests indicate that the estimated models have high validity. The Sargan and Hansen statistics are insignificant in all three models (p-values greater than 0.05), indicating the validity of the instruments used and the satisfaction of the over-identifying restriction conditions. This finding suggests that the selected instruments are appropriate for controlling endogeneity and that the estimation results are reliable. The Arellano–Bond autocorrelation tests also yield favorable results. The presence of statistically significant first-order autocorrelation (AR(1)) and the absence of significant second-order autocorrelation (AR(2)) in the error terms confirm the correct specification of the dynamic panel model. The White and Breusch–Pagan tests confirm the presence of heteroskedasticity, which justifies the use of the GMM estimator with a robust covariance matrix. The Durbin–Wu–Hausman endogeneity test is significant in all three models, confirming the necessity of using the instrumental variables approach to control for endogeneity.

Table 6. GMM Estimation Results (Models 1, 2, and 3)

Variable	Model (1): Financial Factors	Model (2): Financial + Behavioral	Model (3): With Interaction Effects
	Coefficient (SE)	Coefficient (SE)	Coefficient (SE)
Lagged dependent variable			
R(-1)	0.187*** (0.034)	0.173*** (0.032)	0.165*** (0.031)
Earnings quality factors			
AQ	0.315*** (0.056)	0.298*** (0.053)	0.276*** (0.051)
EP	0.271*** (0.49)	0.254*** (0.047)	0.239*** (0.045)
EPR	0.203*** (0.043)	0.189*** (0.041)	0.176*** (0.039)
ECON	0.179*** (0.046)	0.167*** (0.044)	0.154*** (0.042)
ESMO	-0.148*** (0.039)	-0.136*** (0.037)	-0.124*** (0.035)
Profitability/financial factors			
EPS	0.238*** (0.045)	0.221*** (0.043)	0.206*** (0.041)
DPS	0.172*** (0.041)	0.159*** (0.039)	0.147*** (0.037)
DY	0.149*** (0.038)	0.138*** (0.036)	0.128*** (0.034)
DK	0.095** (0.031)	0.087** (0.029)	0.081** (0.028)
TAXAVO	-0.192*** (0.044)	-0.176*** (0.042)	-0.163*** (0.040)
ETRV	-0.112** (0.036)	-0.103** (0.034)	-0.095** (0.032)
Growth and performance factors			
GO	0.202*** (0.046)	0.187*** (0.044)	0.173*** (0.042)
AGR	0.214*** (0.048)	0.198*** (0.046)	0.184*** (0.044)
SGR	0.281*** (0.053)	0.259*** (0.050)	0.241*** (0.048)
ROA	0.351*** (0.065)	0.323*** (0.062)	0.298*** (0.059)
ROE	0.428*** (0.069)	0.394*** (0.066)	0.364*** (0.063)
ROI	0.261*** (0.051)	0.241*** (0.048)	0.223*** (0.046)
Control variables			
SIZE	0.091** (0.031)	0.084** (0.029)	0.078** (0.028)
FIRMAGE	-0.068* (0.028)	-0.062* (0.027)	-0.057* (0.025)
Behavioral factors			
RSI	–	0.241*** (0.050)	0.187*** (0.045)
PSY	–	0.196*** (0.044)	0.152*** (0.040)
ATR	–	0.284*** (0.057)	0.219*** (0.051)
BSI	–	0.203*** (0.046)	0.158*** (0.042)

FRIMAGE	–	0.178** (0.042)	0.138** (0.038)
Interaction effects			
AQ × RSI	–	–	0.146*** (0.039)
EP × RSI	–	–	0.122*** (0.035)
AQ × ATR	–	–	0.172*** (0.043)
EP × ATR	–	–	0.138*** (0.037)
ROE × PSY	–	–	0.161*** (0.042)
DPS × BSI	–	–	0.132*** (0.036)
TAXAVO × ATR	–	–	-0.096** (0.032)
SGR × RSI	–	–	0.149*** (0.040)
ROA × FRIMAGE	–	–	0.127*** (0.035)
GO × PSY	–	–	0.118** (0.034)
Model statistics			
Intercept	-0.241 (0.162)	-0.198 (0.151)	-0.173 (0.145)
R-squared	0.4736	0.5247	0.5692
Adjusted R²	0.4629	0.5132	0.5561
F-statistic	44.23***	52.67***	61.45***
RMSE	0.1882	0.1629	0.1487
MAE	0.1456	0.1267	0.1154
Observations	2,193	2,193	2,193
Firms	129	129	129

The results of Model 1 show that various dimensions of earnings quality have a significant effect on future stock returns. Accrual quality (0.315), earnings persistence (0.271), earnings predictability (0.203), and earnings conservatism (0.179) all have positive and statistically significant coefficients at the 1% level, confirming the first hypothesis. This indicates that investors pay attention to earnings quality, and companies with higher-quality earnings experience better future returns. Interestingly, earnings smoothing has a significantly negative coefficient (-0.148), indicating that the market interprets this behavior as a negative signal consistent with earnings manipulation.

In Model 2, after adding behavioral variables, substantial changes are observed. Behavioral indicators such as investor emotions (0.284), Relative Strength Index (0.241), and trade imbalance (0.203) all have positive and significant effects on future returns. The increase in R^2 from 0.4736 in Model 1 to 0.5247 in Model 2 and the decrease in RMSE from 0.1882 to 0.1629 show that adding behavioral variables significantly improves predictive accuracy, confirming the fifth hypothesis.

Model 3, which incorporates interaction effects, exhibits the best performance with an R^2 of 0.5692 and an RMSE of 0.1487. The significance of interaction coefficients such as AQ × RSI (0.146) and EP × ATR (0.138) demonstrates that behavioral biases have a moderating effect on the relationship between financial factors and future returns, which confirms the sixth hypothesis.

Table 7. Diebold–Mariano Test for Comparing Predictive Accuracy of Models

Model comparison	DM statistic	P-value	Result
Model (1) vs. Model (2)	-4.892***	0.0000	Model (2) superior
Model (1) vs. Model (3)	-6.734***	0.0000	Model (3) superior
Model (2) vs. Model (3)	-3.456***	0.0005	Model (3) superior
Comparative metrics			
ΔRMSE (Model 2 – Model 1)	-0.0253	-13.44%	
ΔRMSE (Model 3 – Model 1)	-0.0395	-20.99%	
ΔRMSE (Model 3 – Model 2)	-0.0142	-8.72%	
ΔR² (Model 2 – Model 1)	+0.0511	+10.79%	
ΔR² (Model 3 – Model 1)	+0.0956	+20.19%	
ΔR² (Model 3 – Model 2)	+0.0445	+8.48%	

The Diebold–Mariano test results decisively indicate the superiority of the enhanced models. The DM statistic for comparing Models 1 and 2 is -4.892 with a p-value less than 0.0001, showing that Model 2 (including behavioral variables) is significantly superior to Model 1 (financial variables only). This superiority is confirmed by a 13.44% reduction in RMSE and a 10.79% increase in R^2.

The comparison of Models 1 and 3 also shows a significant difference (DM = -6.734, $p < 0.0001$). Model 3, which includes interaction effects, reduces RMSE by 20.99% and increases R^2 by 20.19%. Even the comparison between Models 2 and 3 indicates that adding interaction terms generates meaningful improvement (DM = -3.456 and RMSE reduction of 8.72%). These results clearly highlight the importance of simultaneously incorporating financial factors, behavioral factors, and their interaction effects when predicting future stock returns.

Table 8. Results of Artificial Neural Network Simulation – Stage 1 (Without Behavioral Variables)

Variable	Normalized importance (%)	Rank	Raw importance	Cumulative impact (%)
ROE	100.00	1	0.2847	18.73
AQ	87.34	2	0.2486	35.08
EP	76.45	3	0.2177	49.39
SGR	68.92	4	0.1962	62.29
ROA	64.28	5	0.1830	74.32
EPS	58.73	6	0.1672	85.31
GO	53.41	7	0.1520	95.31
AGR	48.16	8	0.1371	104.32
EPR	44.89	9	0.1278	112.73
ROI	41.23	10	0.1174	120.45
DPS	37.56	11	0.1069	127.48
ECON	34.12	12	0.0971	133.87
DY	30.87	13	0.0879	139.65
TAXAVO	27.94	14	0.0796	144.88
SIZE	24.78	15	0.0706	149.52
DK	21.45	16	0.0611	153.54
ESMO	18.92	17	0.0539	157.08
ETRV	16.34	18	0.0465	160.14
FIRMAGE	13.67	19	0.0389	162.70
Network performance metrics				
MSE (Training)	0.0241			
MSE (Validation)	0.0278			
MSE (Testing)	0.0293			
R^2 (Training)	0.8267			
R^2 (Validation)	0.8034			
R^2 (Testing)	0.7918			
MAPE (%)	12.34			
Accuracy (Direction)	73.45%			

The results of the first stage of ANN simulation, which considers only financial variables, show that return on equity (ROE), with a normalized importance of 100%, is the most important variable in rational investor decision-making. Accrual quality (87.34%) and earnings persistence (76.45%) rank second and third, confirming the importance of earnings quality. Interestingly, the first 18 variables account for about 162.73% of the cumulative impact, indicating the dominance of a limited number of key variables in decision-making. The network performance metrics are also satisfactory. R^2 for the training, validation, and testing datasets are 0.8267, 0.8034, and 0.7918, respectively, indicating good predictive power and the absence of severe overfitting. A MAPE of 12.34% and directional accuracy of 73.45% also indicate acceptable performance in predicting future returns. These results provide a solid baseline for comparison with the second stage (with behavioral variables).

Table 9. Results of Artificial Neural Network Simulation – Stage 2 (With Behavioral Variables)

Variable / metric	Normalized importance (%)	Rank	Rank change	Raw importance	Cumulative impact (%)
ATR	100.00	1	NEW	0.3124	17.89
ROE	94.67	2	↓1	0.2957	34.83
AQ	82.45	3	↓1	0.2576	49.59
RSI	76.89	4	NEW	0.2401	63.35
EP	71.23	5	↓2	0.2225	76.09
SGR	65.34	6	↓2	0.2041	87.79
BSI	59.78	7	NEW	0.1867	98.49
ROA	54.12	8	↓3	0.1690	108.17
PSY	49.67	9	NEW	0.1551	117.06
FRIMAGE	45.23	10	NEW	0.1413	125.15
EPS	41.89	11	↓5	0.1308	132.64
GO	38.45	12	↓5	0.1201	139.52
AGR	35.67	13	↓5	0.1114	145.90
EPR	32.78	14	↓5	0.1024	151.77
ROI	29.91	15	↓5	0.0934	157.12
DPS	27.34	16	↓5	0.0854	162.01
ECON	24.89	17	↓5	0.0778	166.47
DY	22.56	18	↓5	0.0705	170.51
TAXAVO	20.45	19	↓5	0.0639	174.17
SIZE	18.23	20	↓5	0.0570	177.44
DK	15.89	21	↓5	0.0496	180.28
ESMO	13.67	22	↓5	0.0427	182.73
ETRV	11.45	23	↓5	0.0358	184.78
FIRMAGE	9.34	24	↓5	0.0292	186.45
Network performance metrics					
MSE (Training)	0.0189				
MSE (Validation)	0.0224				
MSE (Testing)	0.0237				
R^2 (Training)	0.8389				
R^2 (Validation)	0.8291				
R^2 (Testing)	0.8643				
MAPE (%)	9.78				
Accuracy (Direction)	78.92%				
Improvement vs. Stage 1 (ΔMSE)	+21.6% (MSE)			0.0189	

With the inclusion of behavioral variables, substantial changes in the relative importance of variables are observed. Investor emotions (ATR), with a normalized importance of 100%, rises to first place and becomes the most influential variable affecting future returns. ROE falls to second place (94.67%), and accrual quality to third place (82.45%). The entry of four new behavioral variables (ATR, RSI, BSI, PSY, and FRIMAGE) into the top 10 variables indicates the high importance of psychological factors in investors' decision-making. A notable point is the decline in rank of all financial variables. For example, earnings per share drops from rank 6 to 11, growth opportunities from 7 to 12, and asset growth from 8 to 13. These changes show that in the presence of behavioral information, the relative weight of fundamental factors in decision-making decreases. The improvement in performance metrics is also noteworthy: R^2 increases from 0.7918 to 0.8291, MAPE decreases from 12.34% to 9.78%, and directional accuracy increases from 73.45% to 78.92%. A 21.6% reduction in MSE compared to Stage 1 provides strong support for the fifth hypothesis of the study.

Table 10. Friedman Test for Changes in Variable Ranks (Hypothesis 6)

Variable	Rank – Stage 1	Rank – Stage 2	Rank change	Mean rank	Friedman rank
ROE	1	2	-1	1.50	1
AQ	2	3	-1	2.50	2
EP	3	5	-2	4.00	4
SGR	4	6	-2	5.00	5
ROA	5	8	-3	6.50	7
EPS	6	11	-5	8.50	10
GO	7	12	-5	9.50	11
AGR	8	13	-5	10.50	12
EPR	9	14	-5	11.50	13
ROI	10	15	-5	12.50	14
ATR	–	1	NEW	1.00	–
RSI	–	4	NEW	4.00	–
BSI	–	7	NEW	7.00	–
PSY	–	9	NEW	9.00	–
FRIMAGE	–	10	NEW	10.00	–
Friedman test statistic	$\chi^2 = 187.45$***				
Degrees of freedom	df = 18				
P-value	$p < 0.0001$				
Kendall’s W	0.682				
Wilcoxon signed-rank test	Z = -8.923***				
P-value (Wilcoxon)	$p < 0.0001$				
Source: Research calculations with SPSS 26					

The Friedman test, with a test statistic of $\chi^2 = 187.45$ and a p-value less than 0.0001, shows significant changes in the ranking of variables before and after the inclusion of behavioral factors. This finding decisively confirms the sixth hypothesis of the study regarding the moderating role of behavioral biases. Kendall’s W of 0.682 also indicates a strong and significant association between the rankings. A closer analysis of the changes reveals that ROE and AQ, with a one-step decline, have experienced the smallest change, whereas variables such as EPS, GO, AGR, EPR, and ROI have each dropped by five positions. The Wilcoxon paired-samples test ($Z = -8.923$, $p < 0.0001$) also considers these changes significant. These findings indicate that the presence of behavioral factors not only alters the absolute importance of variables but also transforms their relative importance structure, highlighting the complexity of investors’ decision-making processes in the presence of multidimensional information.

Table 11. SHAP Analysis for Model Interpretability

Variable	Mean SHAP Value	SHAP Value Standard Deviation	Relative Contribution (%)	Importance Rank
ATR	0.0847	0.0623	14.52	1
ROE	0.0791	0.0589	13.55	2
AQ	0.0668	0.0512	11.44	3
RSI	0.0624	0.0487	10.69	4
EP	0.0578	0.0445	9.90	5
SGR	0.0531	0.0421	9.10	6
BSI	0.0485	0.0398	8.31	7
ROA	0.0439	0.0367	7.52	8
PSY	0.0403	0.0341	6.90	9
FRIMAGE	0.0367	0.0318	6.29	10
Top 5 Interaction Effects				
AQ × ATR	0.0289	0.0234	4.95	–
ROE × RSI	0.0267	0.0219	4.58	–
EP × ATR	0.0245	0.0201	4.20	–
SGR × RSI	0.0234	0.0189	4.01	–
ROA × FRIMAGE	0.0223	0.0178	3.82	–

SHAP Correlation Metrics	
Correlation with model output	r = 0.9523***
R^2 (SHAP explanation)	0.9069
Mean absolute SHAP	0.0421

The SHAP analysis, which is an advanced method for interpreting machine learning models, provides reliable results. Investor emotion (ATR), with an average SHAP value of 0.0847 and a relative contribution of 14.52%, is the most influential variable. ROE (13.55%) and accrual quality (11.44%) follow, confirming their importance. The high variability of SHAP values (as seen in the standard deviation) indicates that the impact of each variable differs depending on conditions and interactions with other variables.

The analysis of major interaction effects also provides valuable insights. The interaction of AQ × ATR, with a contribution of 4.95%, is the most significant interaction, showing that the effect of accrual quality on returns strongly depends on market sentiment. Similarly, ROE × RSI (4.58%) and EP × ATR (4.20%) represent meaningful interactions. The high correlation (0.9523) between SHAP values and the model output, along with R^2 = 0.9069, demonstrates that SHAP explains model behavior very well and is highly reliable.

Table 12. Results of Regression Decision Tree Algorithm

Evaluation Metric	Value	Performance Indicator
Accuracy Metrics		
R^2 (Training)	0.8834	Excellent
R^2 (Testing)	0.7645	Good
MSE (Training)	0.0167	–
MSE (Testing)	0.0334	–
RMSE (Testing)	0.1827	–
MAE (Testing)	0.1412	–
MAPE (Testing)	11.23%	–
Model Parameters		
Maximum tree depth	12	–
Minimum samples for split	15	–
Minimum samples in leaf	8	–
Total number of nodes	487	–
Number of leaves	244	–
Number of variables used	24	All
Variable Importance (Gini Importance)		
ATR	0.1834	Rank 1
ROE	0.1567	Rank 2
AQ	0.1245	Rank 3
RSI	0.0989	Rank 4
EP	0.0823	Rank 5
SGR	0.0756	Rank 6
BSI	0.0678	Rank 7
ROA	0.0612	Rank 8
PSY	0.0534	Rank 9
FRIMAGE	0.0489	Rank 10
Top 3 Decision Rules		
Rule 1	IF ATR > 0.45 AND ROE > 0.35 THEN R = High	Confidence: 0.89
Rule 2	IF AQ > 0.40 AND RSI > 0.55 THEN R = High	Confidence: 0.85
Rule 3	IF ATR < 0.25 AND TAXAVO > 0.15 THEN R = Low	Confidence: 0.82

The regression decision tree shows strong performance with $R^2 = 0.8834$ for training data and 0.7645 for testing data. RMSE = 0.1827 and MAPE = 11.23% are also acceptable. The model parameters—maximum depth of 12, 487 nodes, and 244 leaves—indicate an appropriate complexity level without extreme overfitting.

Variable importance based on Gini importance aligns with SHAP results: ATR (0.1834), ROE (0.1567), and AQ (0.1245) are the most important variables. The extracted decision rules are also highly informative. Rule 1 indicates with 89% confidence that if ATR > 0.45 and ROE > 0.35, the return will be high. Rule 2 shows that AQ > 0.40 combined with RSI > 0.55 also leads to high returns. These rules can be used as practical guidelines for investment decision-making.

Table 13. Results of LSTM Models for Time-Series Prediction

Evaluation Metric	Simple LSTM	Bidirectional LSTM	Multilayer LSTM	Stacked LSTM
Accuracy Metrics				
R^2 (Testing)	0.7234	0.7589	0.7912	0.8145
RMSE	0.1634	0.1523	0.1418	0.1336
MAE	0.1289	0.1201	0.1123	0.1067
MAPE (%)	10.45	9.78	9.12	8.67
Directional Accuracy (%)	74.56	76.89	79.34	81.23
Convergence Metrics				
Epochs to convergence	78	92	115	134
Training time (min)	12.3	18.7	26.4	35.8
Best validation loss	0.0267	0.0234	0.0201	0.0178
Time-series Feature Importance				
t–1 (previous)	0.342	0.356	0.378	0.391
t–2	0.278	0.289	0.301	0.315
t–3	0.234	0.245	0.256	0.267
t–4	0.146	0.110	0.065	0.027

Comparing the four LSTM architectures shows that the Stacked LSTM model has the best performance, with $R^2 = 0.8145$ and RMSE = 0.1336. It also has the lowest MAE (0.1067) and MAPE (8.67%). A directional accuracy of 81.23% indicates strong ability to correctly predict the direction of price movements. However, the training time of 35.8 minutes and 134 epochs indicates higher computational complexity.

The bidirectional LSTM also performs well with $R^2 = 0.7589$, although weaker than the Stacked LSTM. The simple LSTM, with the fastest training time (12.3 minutes) and the fewest epochs (78), is a good option when speed is a priority. The feature importance analysis shows that the most recent observation (t–1) has the highest importance (0.391 in Stacked LSTM), and importance decreases with longer time lags—which is logically expected.

Table 14. Results of Transformer Models for Return Prediction

Metric	Base Transformer	Transformer + Attention	Multi-head Transformer
Accuracy Metrics			
R^2 (Testing)	0.7891	0.8234	0.8512
RMSE	0.1425	0.1304	0.1196
MAE	0.1134	0.1045	0.0967
MAPE (%)	9.34	8.56	7.89
Directional Accuracy (%)	78.45	81.67	84.23
Model Parameters			
Number of attention heads	4	8	16
Number of encoder layers	2	4	6
Embedding dimension	64	128	256
Feed-forward dimension	256	512	1024
Dropout rate	0.1	0.15	0.2
Attention Analysis			

Attention weight – behavioral variables	0.423	0.456	0.489
Attention weight – financial variables	0.577	0.544	0.511
Attention weight – interactions	–	0.278	0.312
Computational Efficiency			
Training time (min)	28.4	45.7	67.3
Inference time (ms/sample)	3.2	5.8	9.4
Number of parameters (M)	2.1	8.7	24.3

Transformer models, which rely on the attention mechanism, demonstrate superior performance compared to LSTM models. The Multi-head Transformer with 16 attention heads achieves $R^2 = 0.8512$ and RMSE = 0.1196, outperforming even the best LSTM model. A MAPE below 8% and directional accuracy of 84.23% are also notable. This improvement shows that the attention mechanism helps the model learn more complex relationships among variables.

The attention weights offer interesting insights. In the base model, financial variables capture 0.577 of total attention while behavioral variables capture 0.423. However, in the Multi-head Transformer, this becomes 0.511 and 0.489, respectively, and interactions capture 0.312 of the attention. This indicates that the more advanced model recognizes the importance of both variable types and their interactions. Of course, this improved performance comes with higher computational cost: 67.3 minutes of training time and 24.3 million parameters.

Table 15. Monte Carlo Simulation Results for Portfolio Optimization

Scenario	Expected return	Risk (standard deviation)	Sharpe ratio	VaR (95%)	CVaR (95%)	Probability of positive return
Portfolio based on GMM model	0.1847	0.2134	0.865	-0.1523	-0.1876	0.6834
Portfolio based on ANN	0.2134	0.2056	1.038	-0.1389	-0.1712	0.7245
Portfolio based on LSTM	0.2289	0.2012	1.138	-0.1312	-0.1634	0.7456
Portfolio based on Transformer	0.2456	0.1978	1.242	-0.1245	-0.1567	0.7689
Portfolio based on game theory + Bayes	0.2612	0.1934	1.351	-0.1178	-0.1489	0.7923
Market index (Benchmark)	0.1423	0.2467	0.577	-0.1923	-0.2345	0.6234
Relative performance metrics						
Jensen’s alpha (game model)	0.1189***	–	–	–	–	–
Treynor ratio (game model)	0.2134	–	–	–	–	–
Information ratio (game model)	1.567	–	–	–	–	–
Maximum drawdown (game model)	-0.1567	–	–	–	–	–
Calmar ratio (game model)	1.667	–	–	–	–	–
Simulation parameters						
Number of simulations	100,000	–	–	–	–	–
Forecast horizon (months)	12	–	–	–	–	–
Confidence level	95%	–	–	–	–	–
Seed (reproducibility)	42	–	–	–	–	–

The Monte Carlo simulation results with 100,000 iterations clearly demonstrate the superiority of the model based on game theory and Bayes’ rule. This portfolio achieves an expected return of 0.2612, a risk of 0.1934, and a Sharpe ratio of 1.351, which is significantly better than the market index (return of 0.1423 and Sharpe ratio of 0.577) and the other models. Even compared to the Transformer model, which has the second-best performance, a substantial improvement is observed (around 6% in return).

The risk metrics also indicate a favorable situation. VaR at the 95% level is -0.1178 and CVaR is -0.1489, which are considerably better than the market index (respectively -0.1923 and -0.2345). The probability of a positive return, 0.7923, also reflects high stability and reliability of this strategy. The significant Jensen's alpha (0.1189) shows that this portfolio generates a statistically meaningful excess return relative to the market. The high Calmar ratio of 1.667 further indicates an appropriate return adjusted for maximum drawdown, which decisively confirms the seventh research hypothesis.

Table 16. Portfolio Optimization Results Based on Game Theory and Bayes' Rule

Firm	Portfolio weight (%)	Probability of buy position	Probability of sell position	Quality score	Predicted return	Rank
Firm A	18.45	0.8734	0.1266	0.9123	0.3456	1
Firm B	16.23	0.8512	0.1488	0.8867	0.3212	2
Firm C	14.78	0.8345	0.1655	0.8634	0.3089	3
Firm D	12.89	0.8123	0.1877	0.8412	0.2934	4
Firm E	11.34	0.7967	0.2033	0.8234	0.2812	5
Firm F	9.67	0.7756	0.2244	0.8045	0.2678	6
Firm G	8.23	0.7534	0.2466	0.7834	0.2534	7
Firm H	6.89	0.7312	0.2688	0.7623	0.2412	8
Firm I	5.45	0.7089	0.2911	0.7412	0.2289	9
Firm J	3.12	0.6834	0.3166	0.7189	0.2145	10
Others (19 firms)	-7.05	–	–	–	–	–

The optimization results show that the top 10 firms constitute about 105% of the portfolio (the deviation from 100% is due to 7.05% short-selling in other firms). Firm A, with a weight of 18.45%, has the largest share because of its high probability of a buy position (0.8734), excellent quality score (0.9123), and high predicted return (0.3456). The weights are distributed in a descending pattern from 18.45% to 3.12%, indicating a focus on higher-quality firms.

Interestingly, there is a strong correlation between the probability of a buy position, the quality score, and the predicted return. Firms with higher ranks in all three criteria receive larger weights. For example, Firm D, with a buy probability of 0.8123, quality score of 0.8412, and predicted return of 0.2934, has a weight of 12.89%, whereas Firm J, with a buy probability of 0.6834, quality score of 0.7189, and return of 0.2145, has only a 3.12% weight. This pattern shows that the optimization model properly incorporates quality and return potential in capital allocation.

Table 17. Summary of Research Hypothesis Test Results

Hypothesis	Hypothesis content	Test method	Key evidence from findings	Result
Hypothesis 1	Different dimensions of earnings quality have a significant and positive effect on future stock returns.	Generalized method of moments (GMM) regression	Positive and significant coefficients for AQ, EP, EPR, ECON, and a negative coefficient for ESMO in Model (1) (Table 6).	Confirmed
Hypothesis 2	Dividend policy and tax strategies have a significant effect on future returns.	Generalized method of moments (GMM) regression	Positive and significant coefficients for DPS, DY, DK and negative and significant coefficients for TAXAVO, ETRV in Model (1) (Table 6).	Confirmed
Hypothesis 3	Firm growth and performance factors have a significant and positive effect on future returns.	Generalized method of moments (GMM) regression	Positive and significant coefficients for GO, AGR, SGR, ROA, ROE, ROI in Model (1) (Table 6).	Confirmed
Hypothesis 4	Firm characteristics (control variables) have a significant effect on the relationship between earnings quality and future returns.	Generalized method of moments (GMM) regression	Positive and significant coefficient for SIZE and negative and significant coefficient for FIRMAGE in Model (1) (Table 6).	Confirmed
Hypothesis 5	Adding behavioral biases increases the predictive accuracy of future returns.	1. Comparison of GMM models 2. Diebold–Mariano test 3. Comparison of ANN accuracy	Increase in R^2 and decrease in RMSE from Model (1) to Models (2) and (3) (Table 6). Significant DM statistic and improved ΔRMSE and ΔR^2 (Table 7). Improved ANN performance metrics in Stage 2 (Table 9).	Confirmed

Hypothesis 6	Behavioral biases have a moderating role in the relationship between financial factors and future returns.	1. GMM model with interaction effects 2. Friedman test for changes in variable rankings in ANN	Significant interaction coefficients (λ) in Model (3) (Table 6). Significant Friedman statistic (χ^2 = 187.45) and Wilcoxon statistic (Z = -8.923), indicating changes in the ranking of financial variables after inclusion of behavioral variables (Table 10).	Confirmed
Hypothesis 7	Simulation of rational decision-making can increase the return of the optimized portfolio significantly above the market.	Monte Carlo simulation for portfolio optimization	Expected return and Sharpe ratio of the portfolio based on the game- and Bayes-based model (0.2612 and 1.351, respectively) are significantly higher than the market index (0.1423 and 0.577) and other models (Table 15).	Confirmed

Table 17, as a comprehensive summary table, provides a complete picture of the status of all seven research hypotheses. It shows that all hypotheses have been tested and fully confirmed using a wide range of complementary statistical and computational methods. Hypotheses 1 to 4 were tested using generalized method of moments (GMM) regression, which is capable of controlling endogeneity, heteroskedasticity, and autocorrelation problems. Statistical significance of coefficients at the 1% and 5% levels for various dimensions of earnings quality (AQ, EP, EPR, ECON, and ESMO), dividend policy and tax strategies (DPS, DY, DK, TAXAVO, and ETRV), growth and performance factors (GO, AGR, SGR, ROA, ROE, and ROI), and control variables (SIZE and FIRMAGE) is documented in Table 6.

Hypothesis 5, which examines the impact of adding behavioral biases on predictive accuracy, was confirmed through three complementary approaches: first, comparison of GMM models (Model 1 versus Model 2), which showed an increase in R^2 from 0.4736 to 0.5247 and a decrease in RMSE from 0.1882 to 0.1629; second, the Diebold–Mariano test, which confirmed the significant superiority of Model 2 with a DM statistic of -4.892 and a p-value less than 0.0001; and third, improvement in artificial neural network performance metrics in Stage 2 (Table 9), with an increase in R^2 from 0.7918 to 0.8291 and a 21.6% reduction in MSE.

Hypothesis 6, the most complex hypothesis, addressing the moderating role of behavioral biases, was also confirmed from two angles: first, the significance of interaction coefficients (λ) in Model 3 (Table 6), indicating meaningful cross-effects between financial and behavioral variables; and second, the nonparametric Friedman and Wilcoxon tests, which confirmed significant changes in variable rankings after the inclusion of behavioral factors (χ^2 = 187.45, Z = -8.923, both with $p < 0.0001$).

Hypothesis 7, which concerns the ability of rational decision-making simulation to generate superior returns, was tested through Monte Carlo simulation with 100,000 iterations over a 12-month horizon. The results in Table 15 show that the optimized portfolio based on game theory and Bayes' rule produces an expected return of 0.2612, a Sharpe ratio of 1.351, and a significant Jensen's alpha of 0.1189, which are substantially higher than the market index (return 0.1423 and Sharpe ratio 0.577) and other competing models. This table also shows that methodological diversity (combining econometrics, machine learning, simulation, and statistical tests) not only enhances the robustness of the results but also allows triangulation and cross-validation of findings, thereby markedly strengthening confidence in the research outcomes.

Discussion and Conclusion

The results of this study provide strong empirical evidence that multidimensional earnings quality exerts a significant and positive influence on future stock returns when evaluated through dynamic panel GMM estimation, neural-network–based simulation, and deep-learning architectures. The finding that accrual quality, earnings persistence, earnings predictability, and accounting conservatism each display positive and statistically strong coefficients aligns with the long-standing premise that high-quality earnings enhance the informational usefulness of financial statements for investors (7). This evidence is consistent with research emphasizing that faithful representation and sustainable earnings streams enable investors to revise expectations about future performance more accurately and therefore demand lower risk premiums (1). In the Tehran Stock Exchange (TSE), where informational frictions and disclosure asymmetries are more pronounced, the finding that earnings quality dimensions strongly contribute to future returns further supports the argument that high-quality reporting mitigates agency problems and improves the efficiency of capital allocation (8). The negative relationship between earnings smoothing and future returns also reinforces the notion that income-smoothing is viewed as a low-quality signal, often interpreted by investors as indicative of opportunistic earnings management rather than stability (10).

The introduction of behavioral variables into the second model substantially increased explanatory power, as reflected in the rise in R^2 and the reduction in RMSE. Behavioral indicators such as investor emotions (ATR), relative strength index, psychological-line metrics, and trading imbalance demonstrated statistically strong and positive effects. This finding is consistent with research highlighting that investor sentiment meaningfully influences return formation in the TSE, affecting discount rates, cash flows, and realized performance in ways not fully explicable by fundamentals alone (25). Similarly, research grounded in behavioral finance suggests that market-wide psychological factors cause deviation from purely rational Bayesian updating, generating systematic patterns in return predictability (6). The strong influence of ATR, in particular, matches the findings of studies that incorporate sentiment-based predictors in advanced forecasting models, such as those analyzing P/E ratios with LSTM and panel regression structures (20). The improved accuracy metrics in Model 2 confirm that a behavioral–fundamental hybrid structure more closely approximates how investors in emerging markets actually process information.

When interaction terms were incorporated in the third model, the results indicated that behavioral factors significantly moderate the relationship between earnings quality variables and future returns. Interaction effects such as AQ × ATR, EP × ATR, ROE × RSI, and SGR × RSI showed positive and significant coefficients, meaning that the association between fundamentals and returns strengthens when market sentiment is favorable. This finding aligns closely with the idea that accounting signals are not evaluated in isolation, but rather interpreted within a broader behavioral and institutional context. Similar interaction patterns are reported in studies that show how real earnings management, governance mechanisms, and dividend policy jointly shape market responses to earnings news (10, 11). The moderation results also support research demonstrating that reporting quality affects investment efficiency more strongly when ownership structure, financial constraints, or behavioral biases shape how investors interpret accounting information (12, 13). Taken together, these findings suggest that earnings quality should be understood not only as a static characteristic of financial reporting but also as a dynamic signal whose effectiveness depends on the behavioral state of the market.

The machine-learning component of this study provides further evidence that both fundamental and behavioral variables are necessary for accurately forecasting stock returns. In the artificial neural network simulation, earnings-related variables such as ROE, accrual quality, and earnings persistence ranked among the most important drivers of rational decision-making in the absence of behavioral inputs. However, once behavioral indicators were introduced, ATR, RSI, psychological-line values, and FRIMAGE entered the top tier and significantly displaced the dominance of purely financial predictors. This drop in relative importance mirrors findings from the literature showing that machine-learning models incorporating behavioral and market sentiment variables outperform fundamental-only models in predictive accuracy (18, 19). It also corresponds with evidence from Iranian markets showing that forecasting accuracy improves when non-linear interactions and behavioral components are considered (21, 22).

The SHAP interpretability analysis reinforces these dynamics by quantifying the marginal impact of each predictor. The fact that ATR produced the highest SHAP values highlights the centrality of investor emotions in shaping return expectations in high-volatility environments, which supports international findings on sentiment-driven valuation adjustments (6). The prominent SHAP contributions of AQ, ROE, EP, and SGR align with the classical understanding that quality-driven fundamental signals anchor investor beliefs, even when behavioral forces are present (1, 7). SHAP-based interaction analysis, which revealed the importance of AQ × ATR and ROE × RSI, corroborates the econometric evidence that sentiment reframes how the market interprets earnings information. Similar insights are reflected in mixed-method studies emphasizing that investors' decision-making quality improves when behavioral distortions are recognized and modeled explicitly (5, 26).

The decision-tree regression model further validated these relationships by producing clear, interpretable rules demonstrating that high ATR combined with strong ROE or high AQ combined with elevated RSI predicts favorable future returns. These rules are consistent with empirical research arguing that investment decisions improve when both informational quality and market conditions are integrated into a coherent decision framework (5). Additionally, the dominance of ATR, ROE, and AQ in the Gini-importance rankings aligns with previous findings that fundamental indicators serve as anchors while behavioral factors provide accelerators or amplifiers (24, 25).

The deep-learning results add further granularity. Across LSTM configurations, the stacked LSTM architecture achieved the highest predictive accuracy, consistent with international evidence that multi-layer memory structures handle long-term time dependencies effectively in return series (19). The Transformer models, however, outperformed all LSTM variants, which aligns with broader trends in financial machine-learning research where self-attention mechanisms capture cross-sectional and temporal dependencies more efficiently (2). The superior performance of the Multi-Head Transformer is consistent with the argument that attention architectures are especially well-suited to markets with nonlinear, high-frequency, and behaviorally influenced structures (18). Attention-weight analysis also demonstrates that the advanced Transformer became increasingly sensitive to behavioral variables and interaction effects, providing further support for the behavioral–fundamental hybrid theory of market prediction.

Collectively, these results contribute to three major areas in the literature. First, they reinforce the theoretical position that earnings quality significantly enhances return predictability in emerging markets, consistent with research showing the governance and investment-efficiency benefits of high-quality reporting (8, 12). Second, they support the growing body of evidence that investor sentiment plays a critical role in shaping asset prices and can meaningfully alter the mapping from fundamentals to future returns (6, 25). Third, they extend recent work applying machine learning to return forecast modeling by demonstrating that combining financial, behavioral, and interaction features within deep architectures substantially improves prediction accuracy (18-20).

This study's findings therefore suggest that optimal decision-making for investors cannot rely solely on earnings quality or solely on behavioral cues; rather, the two must be integrated into a coherent, interpretable, and simulation-based analytical framework. This aligns with broader research advocating the use of AI-enhanced decision-making systems grounded in accounting transparency and the mitigation of behavioral distortions (5, 24). The results further indicate that the dynamic structure of the TSE—subject to institutional constraints, financing bottlenecks, and sector-specific dynamics—requires models that combine fundamental, behavioral, and deep-learning insights to reflect actual market behavior (3, 4, 29).

Limitations of this study arise primarily from the structure of available data and methodological boundaries. Although the dataset covers more than two decades of firm-year observations, some behavioral and market-microstructure variables were approximated using proxy indicators rather than direct investor-level data, which may limit the granularity of behavioral interpretation. The simulation models, while comprehensive, rely on parameterizations that may behave differently under structural breaks, regulatory changes, or crisis periods. Furthermore, although the use of SHAP and decision-tree rules enhances interpretability, the deep-learning components remain partially opaque despite explainability tools.

Future research should explore the integration of real-time sentiment data, such as social media analytics, news-based embeddings, and higher-frequency trading signals, to enhance behavioral modeling depth. Extending the game-theoretic simulation to multi-agent environments where institutional investors, retail investors, and algorithmic traders interact could yield a richer understanding of market dynamics. Additional work should also evaluate the out-of-sample robustness of these models in crisis periods, regime shifts, and sector-specific analyses, including commodities, petrochemicals, and technology sectors.

From a practical standpoint, the results of this study highlight the importance of investor education focused on understanding the joint influence of earnings quality and behavioral biases. Portfolio managers can benefit from incorporating hybrid behavioral–fundamental models into their decision-support systems, while regulators may consider policies that enhance transparency, reporting quality, and access to sentiment-based market indicators. Analysts and financial institutions can apply the interpretable machine-learning rules derived from this study to strengthen early-warning systems, improve portfolio optimization, and reduce susceptibility to emotionally driven mispricing.

Acknowledgments

We would like to express our appreciation and gratitude to all those who helped us carrying out this study.

Authors’ Contributions

All authors equally contributed to this study.

Declaration of Interest

The authors of this article declared no conflict of interest.

Ethical Considerations

All ethical principles were adheried in conducting and writing this article.

Transparency of Data

In accordance with the principles of transparency and open research, we declare that all data and materials used in this study are available upon request.

Funding

This research was carried out independently with personal funding and without the financial support of any governmental or private institution or organization.

References

1. Dempster GM, Oliver NT. Financial market pricing of earnings quality: Evidence from a multi-factor return model. Open Journal of Business and Management. 2019;7:312-29. doi: 10.4236/ojbm.2019.71021.
2. Gu S, Kelly B, Xiu D. Empirical asset pricing via machine learning. Review of Financial Studies. 2020;33(5):2223-73. doi: 10.1093/rfs/hhaa009.
3. Ghalibaf Asl H, Nadri M, Ebrahimi Bay Salami G, Fallah Shams M. Pathology of financing the tourism industry in Iran within the capital market framework (case study). Quarterly Journal of Capital Market Analysis. 2022;2(4):1-26.
4. Salgi M, Nazari SM. Pathology of production financing with an emphasis on the debt market in Iran. Asset Management and Financing. 2024;11(4):93-120. doi: 10.22108/amf.2024.137439.1795.
5. Hasan SA, Piri P, Chalaki P. Designing an optimal decision-making model for investors: Integrating artificial intelligence and financial reporting transparency. Journal of Asset Management and Financing. 2026;14(2):1-24. doi: 10.22108/amf.2025.143308.1934.
6. Vazirani A, Sarkar S, Bhattacharjee T, Dwivedi YK, Jack S. Information signals and bias in investment decisions: A meta-analytic comparison of prediction and actual performance of new ventures. Journal of Business Research. 2023;155(2):113424. doi: 10.1016/j.jbusres.2022.113424.
7. Vaezi SA, Benabi Qadim R. Earnings quality: A major challenge in accounting. Studies in Accounting and Auditing. 2021;10(38):23-38.
8. Karimi M, Ishaqzadeh A, Salehi Far M. Examining the relationship between financial reporting quality and companies' investment decisions with an emphasis on the governance role of accounting information. Investment Knowledge. 2022;11(42):27-55. doi: 10.30495/faar.2022.693668.
9. Bahmani M, Pourzarandi MA, Minoyi M. Factors influencing stock return forecasting: Utilizing knowledge domain analysis and Delphi-Fuzzy technique. Karafan Scientific Quarterly. 2022;19(2):503-25.
10. Markonah M, Siladjaja M, Simu N. The impact of real earnings quality on the future market value by moderated by the dividend policy. Management Research Studies Journal. 2020;1(1):57-75. doi: 10.56174/mrsj.v1i1.349.
11. Adeneye Y, Kammoun I. Real earnings management and capital structure: Does environmental, social and governance (ESG) performance matter? Cogent Business & Management. 2022;9(1):2130134. doi: 10.1080/23311975.2022.2130134.
12. Dehghan F, Pourheidari A, Khodamipour A. The impact of management forecast quality on investment efficiency considering the role of ownership structure. Financial Accounting and Auditing Research. 2022;14(53):51-76. doi: 10.30495/faar.2022.691686.
13. Mohtashmian SH, Ghadrati H, Arabzadeh M, Jabari H. Factors affecting company investment efficiency and its measurement. Dynamic Management and Business Analysis. 2025;4(4):1-18. doi: 10.22108/amf.2025.143308.1934.
14. Asadi Loya N, Maleki Choobari M, Khordiar S. The impact of modern management accounting techniques on earnings management tendencies during financial crises. Investment Knowledge. 2025;16(63):143-66. doi: 10.22034/jik.2025.78589.4776.
15. Rostami M, Makian SN. Forecasting stock returns in Tehran Stock Exchange: A comparison of Bayesian, exponential smoothing, and Box-Jenkins approaches. Iranian Economic Research. 2022;27(91):189-221. doi: 10.22054/ijer.2022.59528.957.
16. Rostami M, Makiyan SN. Tehran Stock Exchange Return Forecasting: Comparison of Bayesian, Exponential Smoothing and Box Jenkins Approaches. Iranian Journal of Economic Research. 2022;27(91):189-221. doi: 10.22054/ijer.2022.59528.957.
17. Fadaei E, Zareh Behnamiri MJ. Predicting negative stock price shocks with an emphasis on financial ratios. Financial Accounting and Auditing Research. 2022;14(55):181-203.
18. Agusta S, Rakhman F, Mustakini JH, Wijayana S. Enhancing the accuracy of stock return movement prediction in Indonesia through recent fundamental value incorporation in multilayer perceptron. Asian Journal of Accounting Research. 2024;9(4):358-77. doi: 10.1108/AJAR-01-2024-0006.

19. Htun HH, Biehl M, Petkov N. Forecasting relative returns for S&P 500 stocks using machine learning. Financial Innovation. 2024;10:118. doi: 10.1186/s40854-024-00644-0.

20. Dolaeva A, Beliaeva U, Grigoriev D, Semenov A, Rysz M. Analyzing and forecasting P/E ratios using investor sentiment in panel data regression and LSTM models. International Review of Economics & Finance. 2025;98:103840. doi: 10.1016/j.iref.2025.103840.

21. Karimi Pouya MR, Ghambari M, Jamshidi Navid B, Ismailpour M. Examining the accuracy of learning machines in predicting returns from stock price changes using the Raft model, nearest neighbors, and decision trees. Journal of Financial Engineering and Securities Management. 2019;10(38):215-34.

22. Jafari A, Mansouri-Khah M, PourAghajan A. Forecasting stock returns with an emphasis on the role of financial and regulatory criteria using machine learning methods. Accounting and Auditing Studies. 2023;12(45):125-46. doi: 10.22034/iaas.2023.172687.

23. Mohammadi Ladar M, Dadashi A. Stock return prediction models: Estimating the distribution of total market returns and its volatility based on the Laplace distribution. Journal of Judgment and Decision Making in Accounting. 2024;3(9):51-70.

24. AzarSaeed Y, Rostami S. Artificial intelligence and ethical decision-making in accounting and auditing: An analysis of related challenges. Judgment and Decision Making in Accounting. 2023;2(7):114-87. doi: 10.30495/jdaa.2023.705528.

25. Rahmanian Koshki A, Saadat A. The impact of investor sentiment on returns, cash flows, discount rates, and performance of companies listed on the Tehran Stock Exchange. Quarterly Journal of Financial and Economic Policies. 2023;11(41):43-79.

26. Shah Hosseini N, Yelfani E, Khosrovani A. Designing a model for improving investor decision-making and investment efficiency with an emphasis on financial reporting quality: A mixed approach. Accounting, Financial Affairs, and Computational Intelligence. 2024;2(3):120-33.

27. Alipour S, Malekian E, Fakhari H. A network model for data envelopment analysis to evaluate the informational efficiency of reporting units. Financial Accounting and Auditing Research. 2022;14(54):1-48. doi: 10.30495/faar.2022.693668.

28. Siladjaja M, Jasman J. The role of earnings quality and future returns: An illustrative simulation of rational decision model. Journal of Open Innovation: Technology, Market, and Complexity. 2024;10(1):100191. doi: 10.1016/j.joitmc.2023.100191.

29. Alizadeh Chamazakti M, Fattahabadi M, Mahmoudzadeh M, Ghavidel Doostkoui S. The feasibility or impossibility of predicting stock prices: Evidence from the petro-refining industry. Financial Research. 2024;26(1):81-104. doi: 10.22059/frj.2023.359810.1007467.