International Journal of Advanced and Applied Sciences, 13(5) 2026, Pages: 268-282

Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: https://www.science-gate.com/IJAAS.html

Earnings quality of Vietnamese banks: A nonlinear perspective with generalized additive models

Chau Dinh Linh *

Business Administration Faculty, Ho Chi Minh University of Banking, Ho Chi Minh City, Vietnam

ARTICLE INFO

Article history:
Received 7 January 2026
Received in revised form
13 May 2026
Accepted 28 May 2026

Keywords:
Earnings quality
Nonlinear relationships
Generalized additive models
Corporate governance
Vietnamese commercial banks

ABSTRACT

Earnings quality is important in the banking sector because of the complexity of banking operations and strict regulatory requirements. This study investigates the determinants of earnings quality in Vietnamese listed commercial banks during the period 2018–2024. A multidimensional earnings quality index was developed, and Ordinary Least Squares (OLS), panel regression, and Generalized Additive Models (GAMs) were applied to examine both linear and nonlinear relationships. The results show that deposit-based financial leverage has a positive effect on earnings quality in linear models; however, GAM results indicate that this positive effect weakens at higher levels of leverage. Bank size is negatively associated with earnings quality in linear models, while GAM results reveal a nonlinear relationship in which earnings quality decreases at medium levels of bank size but stabilizes or improves among the largest banks. Board gender diversity also demonstrates a nonlinear effect. Furthermore, GAM findings show that gender diversity moderates the relationship between board size and earnings quality. Larger boards are associated with lower earnings quality when female representation is low, but this negative effect becomes weaker as female representation increases. The GAM model provides greater explanatory power than the linear models (Pseudo R^2 = 0.6042 compared to the maximum OLS R^2 = 0.375), highlighting the importance of nonlinear and interaction effects in explaining earnings quality. These findings provide evidence from Vietnamese banks and suggest that governance structure and board composition should be considered in the evaluation of earnings quality.

1. Introduction

Earnings quality plays a fundamental role in ensuring the reliability and usefulness of financial reporting, as it reflects the extent to which reported earnings faithfully represent a firm's underlying economic performance and provide decision-relevant information to stakeholders (Dechow et al., 2010). High-quality earnings enhance transparency, reduce information asymmetry, and facilitate efficient capital allocation. In contrast, low earnings quality can obscure firm performance, mislead investors, and undermine market confidence. These concerns are particularly relevant in the banking sector, where financial reporting is inherently complex due to extensive use of accrual accounting, significant managerial discretion in loan loss provisioning, and the opacity of risk exposures (Bushman and Williams, 2012).

The importance of earnings quality in banks extends beyond firm-level implications to broader financial system stability. Banks serve as critical financial intermediaries, and unreliable earnings reporting may impair market discipline, distort risk assessment, and contribute to systemic vulnerabilities (Bushman and Williams, 2015). Transparent and credible financial reporting is therefore essential for maintaining investor confidence, supporting effective regulatory oversight, and promoting the stability and efficiency of the banking system (Leuz et al., 2003). These issues are particularly salient in emerging economies, where institutional frameworks, corporate governance mechanisms, and regulatory enforcement may still be evolving.

Vietnam provides a particularly relevant context for examining earnings quality in the banking sector.

* Corresponding Author.
Email Address: linhcd@hub.edu.vn
https://doi.org/10.21833/ijaas.2026.05.025
Corresponding author's ORCID profile:
https://orcid.org/0009-0001-9547-625X

Over the past two decades, the Vietnamese banking system has undergone substantial restructuring, including privatization of state-owned banks, improvements in regulatory supervision, and increased integration into international financial markets. These reforms have contributed to enhanced operational efficiency and financial stability. However, challenges related to corporate governance, ownership concentration, and financial transparency remain, especially among listed commercial banks. The variation in governance structures, ownership types, and operational characteristics across listed Vietnamese banks provides a valuable setting for investigating the determinants of earnings quality in an emerging market context.

Vietnam's banking sector operates under strict prudential supervision. The State Bank of Vietnam (SBV) is the primary regulator and oversees financial reporting, risk management, and compliance with prudential standards. Listed banks are also subject to corporate governance and disclosure requirements applicable to publicly listed firms, which require transparent and audited financial reporting. Ownership in Vietnamese banks is relatively concentrated. Many banks have controlling shareholders, including state ownership and strategic investors. These ownership structures may influence governance effectiveness and financial reporting incentives. In addition, regulatory reforms during the sample period strengthened disclosure and risk management requirements, improving governance standards and transparency. This institutional setting is important for understanding how governance characteristics may influence earnings quality in Vietnamese banks.

Despite the extensive international literature on earnings quality, empirical evidence from Vietnam remains limited. Existing studies in Vietnam have largely focused on earnings management using single-dimension proxies and conventional linear regression approaches (Hung and Van, 2020). While these studies provide important insights, they may not fully capture the multidimensional nature of earnings quality, which encompasses attributes such as persistence, predictability, volatility, and smoothness (Dechow et al., 2010). Moreover, the relationships between governance characteristics and earnings quality may be nonlinear, particularly in emerging banking systems characterized by structural heterogeneity and varying governance practices. However, empirical studies examining these multidimensional and potentially nonlinear relationships in the context of listed Vietnamese commercial banks remain scarce.

This study extends the existing literature in several important ways. First, it provides empirical evidence on earnings quality in listed commercial banks in Vietnam, contributing to the limited body of research on financial reporting quality in emerging banking markets. Second, the study adopts a multidimensional earnings quality index that incorporates key attributes of earnings, allowing for a more comprehensive assessment of financial reporting quality than single-dimension measures. Third, the study applies both conventional panel regression models and Generalized Additive Models (GAMs) to examine the determinants of earnings quality, enabling the analysis to capture both linear and potential nonlinear relationships. By integrating multidimensional measurement and flexible modeling techniques, the study offers a more nuanced understanding of earnings quality dynamics in the Vietnamese banking sector.

2. Literature review and hypothesis

2.1. Earnings quality: Concept and measurement

Earnings quality is a central concept in accounting and finance, reflecting the extent to which reported earnings faithfully represent a firm's underlying economic performance and provide useful information for forecasting future performance and cash flows (Dechow et al., 2010). High-quality earnings enhance transparency and reduce information asymmetry between firms and stakeholders, thereby supporting efficient capital allocation and informed decision-making. In contrast, low-quality earnings may distort firm performance, mislead investors, and weaken market discipline.

Prior literature has identified several key attributes that collectively characterize earnings quality. Dechow and Dichev's (2002) accrual-based framework emphasized the importance of the relationship between accruals and cash flows, highlighting that high-quality earnings are supported by reliable accrual estimation (Barth et al., 2016). Subsequent research has expanded this perspective by examining additional attributes such as persistence, predictability, volatility, and smoothness (Dong et al., 2025; Leuz et al., 2003).

Earnings persistence refers to the extent to which current earnings are sustainable and expected to recur in future periods. Higher persistence indicates greater earnings stability and reliability. Predictability reflects the ability of current earnings to forecast future earnings accurately, with more predictable earnings generally considered higher quality. Earnings volatility captures fluctuations in earnings over time, where excessive volatility may indicate instability or increased uncertainty. Earnings smoothness measures the degree to which earnings vary relative to underlying cash flows. While moderate smoothness may reflect stable business operations, excessive smoothness may signal managerial discretion or earnings management (Leuz et al., 2003).

In the banking sector, the assessment of earnings quality is particularly important due to the complexity of financial intermediation, the extensive use of accrual accounting, and managerial discretion in areas such as loan loss provisioning and asset valuation (Bushman and Williams, 2012). These characteristics make a multidimensional approach incorporating persistence, predictability, volatility, and smoothness especially appropriate for evaluating earnings quality in banking institutions.

2.2. Financial leverage and earnings quality

Financial leverage influences managerial incentives and financial reporting quality through monitoring and risk mechanisms. According to agency theory, higher leverage can enhance creditor monitoring, thereby reducing managerial opportunism and improving earnings quality (Anagnostopoulou and Tsekrekos, 2017). However, excessive leverage may increase financial distress risk, potentially encouraging earnings manipulation to avoid covenant violations (Awuye and Aubert, 2022). In the banking sector, deposits represent the primary form of leverage and are generally considered a stable and reliable funding source (Hanson et al., 2015; Allen et al., 2015). A higher deposit-to-total-assets ratio reflects greater funding stability, improved liquidity conditions, and reduced financial uncertainty, which may contribute to more reliable and higher-quality earnings (Vazquez and Federico, 2015). Based on these arguments, this study proposes the following hypothesis:

H1: Financial leverage, measured by the deposit-to-total-assets ratio, is positively associated with earnings quality.

2.3. Bank size and earnings quality

Bank size influences operational complexity, governance, and financial reporting quality. Larger banks often benefit from economies of scale, diversified operations, and better access to financial resources, which may enhance financial stability and reporting reliability (Bertay et al., 2013).

However, a larger size may also increase organizational complexity, reduce transparency, and create challenges in internal monitoring and control (Alessandrini et al., 2010). Greater complexity can increase information asymmetry and managerial discretion in financial reporting. In addition, large banks may face stronger market expectations, potentially increasing incentives for earnings management (Cheng et al., 2011).

Given the potential negative effects of organizational complexity on financial reporting transparency, this study proposes the following hypothesis:

H2: Bank size is negatively associated with earnings quality.

2.4. Board gender diversity and earnings quality

Corporate governance plays a crucial role in ensuring the integrity and reliability of financial reporting. Board composition, including gender diversity, has received increasing attention in the governance literature. Female directors are often associated with enhanced monitoring effectiveness, greater ethical sensitivity, and more prudent decision-making (Srinidhi et al., 2011).

Gender diverse boards may improve governance by strengthening oversight, reducing managerial opportunism, and enhancing transparency in financial reporting. Prior studies find that firms with higher female board representation tend to exhibit lower levels of earnings management and improved financial reporting quality (Srinidhi et al., 2011; Chang et al., 2025). Diverse boards may also promote more comprehensive discussions and reduce groupthink, thereby improving governance effectiveness. Based on these arguments, this study proposes:

H3: Board gender diversity is positively associated with earnings quality.

2.5. Other corporate governance characteristics and earnings quality

Board size reflects the number of directors responsible for monitoring management. Larger boards may provide diverse expertise and perspectives but may also face coordination challenges that reduce monitoring effectiveness (Coles et al., 2008).

Financial expertise among board members enhances the board's ability to understand financial reports, evaluate accounting decisions, and detect potential earnings manipulation (Xie et al., 2003).

The presence of independent or non-executive directors strengthens board independence and improves oversight, thereby enhancing financial reporting quality (Klein, 2002). These governance characteristics are included as control variables to account for their potential influence on earnings quality.

2.6. Nonlinear relationships and earnings quality

Traditional empirical studies often assume linear relationships between governance characteristics and earnings quality. However, financial and governance relationships may be nonlinear. For example, moderate leverage may improve monitoring, while excessive leverage may increase financial distress risk. Similarly, governance characteristics such as board diversity may exhibit diminishing or threshold effects.

Generalized Additive Models provide a flexible framework for modeling nonlinear relationships by allowing the data to determine functional relationships without imposing strict linear assumptions (Friedman et al., 2000). GAMs have been increasingly applied in financial research to capture complex relationships and improve model flexibility.

By applying GAMs, this study examines both linear and nonlinear relationships between governance characteristics and earnings quality, providing a more comprehensive analysis.

2.7. Research gap and contribution in the Vietnamese context

Research on earnings quality in Vietnam remains relatively limited, particularly in the banking sector. Existing studies primarily focus on earnings management using single-dimensional proxies and conventional econometric approaches (Hung and Van, 2020). While these studies provide valuable insights, they may not fully capture the multidimensional nature of earnings quality.

Furthermore, most existing studies rely on linear econometric models, which may not adequately capture complex and potentially nonlinear relationships between governance characteristics and earnings quality. This limitation is particularly relevant in emerging banking systems such as Vietnam, where governance practices, ownership structures, and institutional environments vary significantly across banks.

This study contributes to the literature in several ways. First, it examines earnings quality in listed commercial banks in Vietnam, providing empirical evidence from an emerging banking market. Second, it adopts a multidimensional earnings quality measure incorporating persistence, predictability, volatility, and smoothness, providing a more comprehensive assessment of financial reporting quality. Third, the study applies both panel regression models and Generalized Additive Models to examine potential nonlinear relationships between governance characteristics and earnings quality.

By integrating multidimensional measurement and flexible modeling techniques, this study contributes to a better understanding of earnings quality determinants in the Vietnamese banking sector.

3. Data and methodology

3.1. Data and sample selection

This study employs a panel dataset of listed Vietnamese commercial banks over the period 2016–2024. The Vietnamese banking sector provides an appropriate context for examining earnings quality due to its regulatory reforms, increasing transparency requirements, and diversity in governance structures.

Financial and governance data are obtained from the FiinProX database and the annual report of each bank, a widely used financial data source in Vietnam. The initial dataset consists of 1,092 bank-year observations. Observations with missing values for key variables are excluded to ensure data reliability.

After data cleaning, 161 bank-year observations are retained for constructing the earnings quality indices. Due to the rolling window methodology used to calculate earnings quality measures, the effective regression sample consists of 159 bank-year observations, where all required variables are available. Each observation represents a unique bank-year pair.

Although raw financial data are available from 2016, the effective analytical period for models using earnings quality indices (EQ_window_5 and EQ_window_3) spans from 2018 to 2024, reflecting the data requirements of the rolling window approach. This panel structure allows the analysis to capture both cross-sectional differences across banks and time-series variations within banks.

3.2. Measurement of earnings quality

Earnings quality is quantified through a robust multidimensional index constructed from four widely recognized attributes: persistence, predictability, volatility, and smoothness. These attributes collectively capture various critical facets of earnings reliability and have been extensively employed and validated in prior academic literature (Dechow et al., 2010).

Persistence assesses the extent to which current earnings are indicative of future earnings, reflecting their sustainability. It is estimated using the following first-order autoregressive model:

$$EARN_{i,t} = \alpha_0 + \alpha_1 EARN_{i,t-1} + \varepsilon_{it} \quad (1)$$

In this model, $EARN_{i,t}$ represents the earnings for bank i at time t, and $EARN_{i,t-1}$ is the earnings in the prior period. The coefficient α_1 directly represents earnings persistence. A higher value of α_1 indicates greater earnings sustainability and, consequently, higher earnings quality.

Predictability is measured as the standard deviation of the residuals (ε_{it}) derived from the persistence regression model described above. These residuals capture the unpredictable component of earnings. A larger standard deviation of these residuals implies lower predictability, indicating greater earnings uncertainty and, by extension, lower earnings quality.

Volatility is computed as the standard deviation of earnings over a defined rolling window period. Elevated earnings volatility signifies reduced stability in reported profits, which is generally associated with lower earnings quality as it introduces greater uncertainty about future performance.

Smoothness is determined by the ratio of the volatility of earnings to the volatility of operating cash flows over a specific rolling window:

$$SMOOTH_i = \frac{\sigma(EARN_{i,t})}{\sigma(CFO_{i,t})} \quad (2)$$

This metric captures the inherent relationship between reported earnings and underlying operating cash flows (Leuz et al., 2003). While some degree of smoothness might naturally occur, an unusually low ratio could suggest efforts to manage or smooth earnings, potentially signaling lower quality.

To facilitate robust statistical comparison across different banks and over the analytical period, each individual component of earnings quality (persistence, predictability, volatility, smoothness) is standardized using cross-sectional Z-scores. This annual standardization process ensures that each component contributes equally to the composite index and that the index is comparable across different years and institutions. The formula for standardization is:

$$Z(X_{i,t}) = \frac{X_{i,t}-\mu_t}{\sigma_t} \quad (3)$$

where, $X_{i,t}$ denotes the value of a specific earnings quality attribute for bank i at time t, μ_t is the cross-sectional mean of that attribute at time t, and σ_t is its cross-sectional standard deviation at time t.

The overall composite earnings quality index ($EQ_{i,t}$) is then constructed as an equally weighted average of these standardized components:

$$EQ_{i,t} = \frac{1}{4}\left[Z\left(Persistence_{i,t}\right) - Z\left(Predictability_{i,t}\right) - Z\left(Volatility_{i,t}\right) - Z(Smoothness_{i,t})\right] \quad (4)$$

In this formulation, higher EQ values indicate higher earnings quality. Predictability, volatility, and smoothness are subtracted because lower values of these attributes generally reflect higher earnings quality, characterized by more stable and reliable earnings.

To ensure robustness, the composite index is calculated using two rolling window periods. The five-year rolling window (EQ_window_5) serves as the primary dependent variable, as it captures more stable and persistent earnings quality. The three-year rolling window (EQ_window_3) is used as a robustness check to examine whether the results remain consistent over a shorter time horizon.

3.3. Independent variables

The selection of independent variables is based on prior literature examining the determinants of earnings quality, including financial structure, corporate governance, and macroeconomic conditions. All continuous variables are standardized using Z-scores prior to estimation to improve comparability across variables and reduce potential multicollinearity.

- Deposit – based financial leverage (DFL)

$$DFL_{i,t} = \frac{Deposit_{i,t}}{Total\ Assets_{i,t}} \quad (5)$$

This measure captures the extent to which banks rely on customer deposits as a primary funding source (de Comères, 2025). Deposits are generally considered more stable and less costly than wholesale or market-based funding, contributing to funding stability and earnings reliability (Allen et al., 2015; Hanson et al., 2015). A higher proportion of deposit funding may reduce liquidity risk and support more stable earnings generation (Vazquez and Federico, 2015).

- Bank size (Bank_size)

$$Bank\_size_{i,t} = \ln\left(Total\ Assets_{i,t}\right) \quad (6)$$

This transformation is commonly used to address skewness in asset distribution and to capture scale-related effects (Dang et al., 2018). Bank size reflects differences in operational complexity, economies of scale, and regulatory oversight, all of which may influence financial reporting quality (Di Martino et al., 2017).

- Board characteristics

Corporate governance variables related to board composition are included due to their important role in financial reporting quality (Tan and Taufiik, 2022).

Board size represents the total number of directors and reflects the board's monitoring capacity and diversity of expertise, although excessively large boards may face coordination challenges (Coles et al., 2008).

Female board ratio measures the proportion of female directors and captures gender diversity, which has been associated with stronger monitoring, improved governance, and enhanced reporting quality (Srinidhi et al., 2011; Chang et al., 2025).

Financial expertise reflects the proportion of directors with financial qualifications or experience, enhancing the board's ability to monitor financial reporting and detect potential manipulation (Xie et al., 2003). Non-executive directors ratio measures the proportion of board members independent from management, contributing to objective oversight and improved governance (Klein, 2002). Independent directors ratio further captures the extent of independent oversight and board effectiveness (Vafeas and Vlittis, 2024).

- Macroeconomic variables

Macroeconomic conditions may influence banking performance and earnings quality. Therefore, two control variables are included.

GDP growth reflects overall economic conditions affecting credit demand, asset quality, and bank profitability (Saeed, 2014).

Inflation captures changes in macroeconomic stability that may affect interest rates, funding costs, and financial reporting outcomes (Thi Tran, 2024; Pham et al., 2023).

Including these macroeconomic variables helps control for external economic influences on earnings quality.

3.4. Model specification

The empirical analysis was implemented in Python using Google Colab. Data preparation and descriptive analyses were conducted with pandas and numpy. OLS regressions with heteroskedasticity-consistent standard errors were estimated using statsmodels. Fixed-effects and random-effects panel models with bank-level clustered standard errors were estimated using the linearmodels package, and model selection between FE and RE relied on the Hausman test. To capture potential non-linearities, a Generalized Additive Model was estimated using pyGAM, and partial dependence plots were used to visualize non-linear effects. All continuous regressors were standardized (z-scores) prior to estimation to improve numerical stability and facilitate coefficient comparability.

- Ordinary least squares regression

The baseline empirical model is specified using Ordinary Least Squares regression, which provides a straightforward assessment of linear relationships:

$$EQ_{it} = \beta_0 + \beta_1 DFL_{it} + \beta_2 Bank\_size_{it} + \beta_3 FEMA_{it} + \beta_4 SIZE_{it} + \beta_5 FIEXP_{it} + \beta_6 NON_{it} + \beta_7 GDP_t + \beta_8 INF_t + \varepsilon_{i,t} \quad (7)$$

where, EQ_{it} represents the earnings quality for bank i at time t, and the independent variables are as defined in Section 3.3. Robust standard errors (specifically, HC1 robust standard errors) are applied to correct for potential heteroskedasticity and to ensure more reliable inference, even if the assumption of homoskedasticity is violated.

- Panel data models

To account for unobserved bank-specific heterogeneity that could bias OLS estimates and to better exploit the panel structure of the data, both Fixed Effects and Random Effects models are estimated.

The Fixed Effects (FE) model controls for all time-invariant unobserved individual effects by incorporating a dummy variable for each bank, or by differencing:

$$EQ_{it} = \beta X_{i,t} + \alpha_i + \varepsilon_{it} \quad (8)$$

where, α_i represents a time-invariant, bank-specific intercept (unobserved heterogeneity).

The Random Effects (RE) model assumes that the unobserved individual effects are uncorrelated with the independent variables and are randomly distributed:

$$EQ_{it} = \beta X_{i,t} + u_i + \varepsilon_{it} \quad (9)$$

where, u_i represents time-invariant, bank-specific random effects. Clustered standard errors at the bank level are employed for both FE and RE models to address potential autocorrelation within banks and heteroskedasticity, ensuring robust inference. The Hausman test is subsequently utilized to formally determine whether the FE or RE model is the more appropriate specification for the data, based on the assumption about the correlation between u_i and $X_{i,t}$.

To ensure reliable statistical inference, this study employs Driscoll–Kraay standard errors in the panel regression models. This approach accounts for heteroskedasticity, serial correlation, and cross-sectional dependence, which are common in panel data involving financial institutions operating in the same economic environment.

- Generalized additive models

To explicitly explore and capture potentially complex nonlinear relationships between the independent variables and earnings quality, Generalized Additive Models are employed. GAMs offer a flexible, semi-parametric framework that extends generalized linear models by allowing the linear predictor to depend on smooth functions of the covariates:

$$EQ_{it} = \alpha + s(SIZE_{it}) + s(FEMA_{it}) + s(FIEXP_{it}) + s(DFL_{it}) + s(GDP_t) + s(INF_t) + s(Bank\_size_{it}) + s(SIZE_{it}, FEMA_{it}) + \varepsilon_{it} \quad (10)$$

To examine whether gender diversity moderates the relationship between board size and earnings quality, the model includes an interaction smooth term $s(SIZE_{it}, FEMA_{it})$. This interaction allows the marginal effect of board size on earnings quality to vary across different levels of female board representation, rather than assuming a constant linear effect. This specification provides a more flexible framework for capturing moderation and nonlinear governance effects.

In this model, s(.) represents smooth, non-parametric functions estimated directly from the data (Friedman et al., 2000). These smooth functions allow the data to determine the shape of the relationship without imposing restrictive linear assumptions. This approach enables the identification of nonlinear, threshold, or moderating effects that conventional linear models may fail to capture. The flexibility of GAMs makes them particularly well-suited for complex financial and governance data, where relationships are often nonlinear and context-dependent.

- Diagnostic tests

Several diagnostic tests are conducted to ensure the validity and robustness of the empirical results. Multicollinearity is assessed using Variance Inflation Factors (VIF). VIF values above 10 generally indicate serious multicollinearity. The results show that all VIF values remain within acceptable ranges after standardization, indicating no severe multicollinearity among the independent variables.

To address potential heteroskedasticity, robust standard errors are applied. Specifically, HC1 robust standard errors are used for OLS models, while clustered standard errors are employed for panel data models. This approach ensures reliable statistical inference even when error variance is not constant. Serial correlation within panel data is addressed by clustering standard errors at the bank level. This method accounts for the potential dependence of observations within the same bank over time. The Hausman test is used to determine the appropriate specification between Fixed Effects and Random Effects models. For OLS models, overall model significance is evaluated using the F-statistic. For GAM models, explained deviance is used as a measure of model fit. These diagnostic procedures help ensure that the empirical results are statistically valid and robust.

4. Results and discussion

4.1. Earnings quality: Construction, distribution, and robustness

This section describes the construction and statistical properties of the earnings quality (EQ) index. The EQ index is developed using a multidimensional framework incorporating persistence, predictability, volatility, and smoothness, consistent with prior literature (Dechow et al., 2010). The index is calculated using rolling windows and standardized annually using cross-sectional Z-scores to ensure comparability across banks and over time.

Although raw financial data are available from 2016, the rolling window methodology requires prior observations, resulting in EQ_window_5 and EQ_window_3 being available from 2018 onward. The dataset used for EQ index construction includes 161 bank-year observations over the period 2018–2024.

Table 1 presents the descriptive statistics for EQ_window_5, which serves as our main dependent variable. As anticipated, due to the annual Z-score standardization, the mean of the EQ index is approximately zero. A standard deviation of 0.6806 indicates substantial variation in earnings quality among the banks within the sample. The wide range, spanning from -3.0582 to 1.3597, highlights significant heterogeneity in earnings quality across individual bank-year observations. This broad distribution, encompassing both negative and positive values, confirms that the measure effectively captures meaningful differences in earnings quality and provides sufficient cross-sectional variation essential for robust regression analysis.

Table 2 presents annual descriptive statistics for EQ_window_5 to assess its temporal stability. Consistent with the standardization procedure, the annual mean remains close to zero in all years. The standard deviation ranges from 0.59 in 2020 to 0.74 in 2022, indicating persistent cross-sectional variation across banks.

This pattern suggests that differences in earnings quality primarily reflect structural differences among banks rather than common time-specific effects. These results support the use of the EQ index as a reliable measure for cross-sectional analysis.

Table 1: Descriptive statistics of earnings quality (EQ_window_5)

Variable	N	Mean	SD	Min	25th percentile	Median	75th percentile	Max
EQ_window_5	161	0.0000	0.6806	-3.0582	-0.1777	0.1661	0.3855	1.3597

EQ is constructed as a composite Z-score index based on persistence, predictability, volatility, and smoothness; Standardization is performed annually

Table 2: Earnings quality (EQ_window_5) by year

Year	Observations	Mean	SD	Min	Max
2018	23	0.0000	0.6860	-2.2342	0.5948
2019	23	0.0000	0.7257	-1.5255	0.8066
2020	23	0.0000	0.5946	-1.2277	0.9861
2021	23	0.0000	0.6439	-2.6029	0.6536
2022	23	0.0000	0.7379	-3.0582	0.7640
2023	23	0.0000	0.7011	-1.7820	1.3597
2024	23	0.0000	0.6726	-1.9745	1.1448

Annual means are approximately zero due to cross-sectional standardization

The EQ index combines standardized measures of persistence, predictability, volatility, and smoothness. Table 3 reveals the correlation reveals strong negative associations between EQ and predictability (−0.882), volatility (−0.826), and smoothness (−0.726), and a positive association with persistence (0.280). These findings indicate that volatility and predictability are the primary contributors to variation in earnings quality, as higher volatility and lower predictability are associated with lower EQ values. This pattern is consistent with theoretical expectations and confirms that the EQ index effectively captures the multidimensional nature of earnings quality.

Table 3: Correlation between earnings quality and its components

Variable	Persistence (Z_PERS)	Predictability (Z_PRED)	Volatility (Z_VOL)	Smoothness (Z_SMOOTH)
EQ index	0.280***	-0.882***	-0.826***	-0.726***

***: $p < 0.001$

To assess robustness, the EQ index is also calculated using a three-year rolling window (EQ_window_3). Table 4 shows a strong positive correlation (0.7397) between EQ_window_5 and

EQ_window_3, indicating consistency across alternative measurement windows. Although shorter windows are more sensitive to short-term fluctuations, the small median absolute difference (0.1323) suggests that the relative ranking of banks remains stable. These results support the robustness of the EQ measure and justify the use of EQ_window_5 as the primary dependent variable.

Table 4: Robustness analysis: Comparison between EQ_window_5 and EQ_window_3

Metric	Value
Observations	161
Correlation coefficient	0.7397
Mean absolute difference	0.2728
Median absolute difference	0.1323
Maximum absolute difference	2.3318

4.2. Descriptive statistics

The final sample consists of a balanced panel of 159 bank-year observations from 23 listed Vietnamese commercial banks over the period 2018–2024. This sample size reflects the rolling-window requirements used to construct the earnings quality indices, which require sufficient historical observations for each bank (Dechow et al., 2010). Table 5 presents descriptive statistics for all variables included in the regression analysis. As described in Section 3.3, all continuous independent variables are standardized using Z-scores to improve comparability and reduce scale-related estimation issues.

Both earnings quality measures (EQ_window_5 and EQ_window_3) have means close to zero and standard deviations of approximately 0.68 and 0.65, respectively, reflecting substantial cross-sectional variation. Governance variables, including board size (SIZE), female representation (FEMA), independent directors (INDD), financial expertise (FIEXP), and non-executive directors (NON), exhibit reasonable variation, indicating diversity in governance structures across banks.

The mean deposit-based leverage (DFL) of 0.664 indicates a strong reliance on customer deposits as a funding source, consistent with prior banking literature (Hanson et al., 2015). Bank Size ranges from 10.658 to 14.812, reflecting significant differences in bank scale (Dang et al., 2018). Macroeconomic variables, including GDP growth and inflation, capture economic conditions during the study period and are commonly used control variables in banking research (Pham et al., 2023). The DUAL variable is excluded from regression analysis due to a lack of variation within the sample.

Table 5: Descriptive statistics of variables

Variable	Mean	SD	Min	Max
EQ_window_5	-0.008	0.681	-3.058	1.360
EQ_window_3	-0.008	0.655	-4.493	1.230
SIZE	7.088	1.034	5.000	11.000
FEMA	0.158	0.049	0.000	0.286
INDD	0.224	0.071	0.000	0.333
FIEXP	0.993	0.032	0.833	1.000
NON	0.865	0.055	0.625	1.000
DFL (deposits/total assets)	0.664	0.101	0.442	0.884
BANK_SIZE_log_assets	12.532	1.014	10.658	14.812
GDP	0.057	0.020	0.026	0.080
INF	0.031	0.006	0.018	0.036

4.3. OLS regression with robust standard errors

This study employs Ordinary Least Squares regression models with HC1 robust standard errors. This approach is adopted to address potential issues of heteroskedasticity and other violations of standard linear regression assumptions, thereby ensuring that the estimated standard errors and associated p-values are reliable for valid statistical inference.

We estimate separate OLS models for both EQ_window_5 (as the primary dependent variable) and EQ_window_3 (for robustness checks). For EQ_window_5, four distinct specifications are developed, progressively incorporating increased analytical flexibility through the inclusion of quadratic terms, interaction effects, and higher-order nonlinear features.

- OLS Results for EQ_window_5

Table 6 presents the OLS regression results for EQ_window_5, utilizing HC1 robust standard errors.

Table 6: OLS results for EQ_window_5

Variable	Model 1	Model 2	Model 3	Model 4
Const	-0.0082 (0.856)	-0.0187 (0.717)	-0.0118 (0.792)	-0.0311 (0.489)
SIZE	-0.0975 (0.173)	-0.1109 (0.134)	-0.1119 (0.160)	-0.2406 (0.017)*
FEMA	0.0456 (0.330)	0.0384 (0.407)	0.0346 (0.457)	0.0475 (0.308)
FIEXP	0.0228 (0.360)	0.0152 (0.505)	0.0210 (0.390)	0.0240 (0.606)
NON	0.0093 (0.835)	0.0135 (0.767)	0.0142 (0.758)	0.0085 (0.855)
INDD	0.0069 (0.893)	0.0051 (0.920)	0.0048 (0.925)	-0.0051 (0.920)
DFL	0.1309 (0.030)*	0.1130 (0.038)*	0.1331 (0.026)*	0.1142 (0.011)**
GDP	-0.0020 (0.971)	-0.0028 (0.961)	-0.0019 (0.972)	0.0027 (0.960)
INF	0.0152 (0.772)	0.0177 (0.749)	0.0172 (0.757)	0.0227 (0.675)
BANK_SIZE_log_assets	-0.2807 (0.000)***	-0.2868 (0.000)***	-0.2834 (0.000)***	-0.2798 (0.001)**
SIZE_sq_std		0.0099 (0.569)		
SIZE_x_FEMA_std			0.0232 (0.803)	
SIZE_cube				0.0199 (0.050)*
DFL×BANK_SIZE_log_assets				0.1351 (0.029)*
Model statistics				
R^2	0.328	0.329	0.328	0.375
Adj. R^2	0.297	0.293	0.292	0.337
F-statistic (p-value)	8.809 (4.59e-09)***	7.762 (1.10e-08)***	7.728 (1.29e-08)***	9.918 (7.69e-12)***
Cond. No.	3.24	5.78	3.27	19.7

***: p<0.001; **: p<0.01; *: p<0.05; p-values in parentheses

Two consistent and highly noteworthy findings emerge from the OLS regressions for EQ_window_5 across all four model specifications. First, BANK_SIZE_log_assets consistently exerts a strong negative and highly statistically significant impact ($p < 0.001$ or $p < 0.01$) on earnings quality. This robust finding suggests that larger bank size is associated with a reduction in earnings quality, potentially attributable to increased organizational complexity, reduced transparency, and more intricate financial operations that may obscure the true economic performance (Alessandrini et al., 2010). Second, DFL (deposit-based financial leverage) consistently exhibits a positive and statistically significant effect (at the 5% level or better) across all models. This indicates that a higher reliance on customer deposits for funding significantly enhances financial stability, which in turn leads to improved earnings quality. Deposits are generally considered a more stable, less costly, and "stickier" funding source compared to other liabilities (Allen et al., 2015; Hanson et al., 2015), and a robust, deposit-based funding structure effectively reduces liquidity risk, thereby promoting more consistent and higher-quality earnings (Vazquez and Federico, 2015).

Model 4, which incorporates advanced interaction and cubic terms, demonstrates a notable enhancement in explanatory power, with the R^2 increasing to 0.375, representing the highest R^2 among the OLS models. This model also provides initial evidence suggestive of nonlinear relationships. Specifically, SIZE (board size) becomes statistically significant and negative, while its cubic term (SIZE_cube) is marginally significant and positive, hinting at a non-monotonic effect. Furthermore, the interaction term DFL×BANK_SIZE_log_assets is significant ($p=0.029$), implying that the effect of deposit-based leverage on earnings quality may not be uniform across different bank sizes, suggesting a more complex, conditional relationship.

- OLS Results for EQ_window_3

Table 7 reports the OLS regression results using EQ_window_3 as the dependent variable. The explanatory power of these models is lower than that of those using EQ_window_5, with R^2 values ranging from 0.219 to 0.226, suggesting that shorter rolling windows capture earnings quality with less stability.

Female board ratio (FEMA) is consistently positive and statistically significant ($p < 0.05$) across all specifications, supporting prior evidence that gender diversity improves governance and financial reporting quality (Srinidhi et al., 2011). Bank Size (BANK_SIZE_log_assets) remains negatively significant ($p < 0.05$), consistent with earlier findings and suggesting that larger banks may face greater complexity that reduces earnings quality.

In contrast, deposit-based leverage (DFL) is not statistically significant in these models. This difference indicates that the positive effect of stable deposit funding on earnings quality may be more evident over longer measurement periods, such as the five-year rolling window.

Table 7: OLS regression results for EQ_window_3

Variable	Model 1	Model 2	Model 3
Const	-0.0079 (0.867)	-0.0385 (0.465)	-0.0169 (0.706)
SIZE	-0.0919 (0.244)	-0.1346 (0.091).	-0.1068 (0.207)
FEMA	0.0793 (0.037)*	0.0757 (0.047)*	0.0802 (0.028)*
FIEXP	-0.0344 (0.207)	-0.0358 (0.212)	-0.0341 (0.203)
NON	-0.0550 (0.144)	-0.0481 (0.205)	-0.0569 (0.139)
INDD	0.0094 (0.854)	0.0056 (0.914)	0.0051 (0.922)
DFL	0.0928 (0.205)	0.0802 (0.278)	0.0929 (0.163)
GDP	-0.0037 (0.941)	-0.0044 (0.929)	-0.0031 (0.950)
INF	0.0096 (0.883)	0.0089 (0.891)	0.0091 (0.887)
BANK_SIZE_log_assets	-0.1586 (0.024)*	-0.1427 (0.044)*	-0.1583 (0.024)*
SIZE_sq_std		0.0307 (0.052).	
SIZE_x_FEMA_std			-0.0101 (0.880)
Model statistics			
R-squared	0.219	0.226	0.219
Adj. R-squared	0.177	0.179	0.172
F-statistic	6.178 (6.93e-07)***	6.551 (7.61e-08)***	6.541 (7.83e-08)***
Cond. No.	3.42	6.01	3.45

***: $p < 0.001$; *: $p < 0.05$; p-values in parentheses

4.4. Panel data regression (FE/RE)

To effectively account for unobserved bank-specific characteristics that remain constant over time and to control for potential endogeneity bias arising from such unobservables, this study employs both Fixed Effects and Random Effects panel regression models, as meticulously detailed in Section 3.4. These models incorporate clustered standard errors at the bank level, a robust methodological choice aimed at addressing potential issues of heteroskedasticity and serial correlation within individual banks, consistent with the rigorous diagnostic procedures outlined in Section 3.4.

- Panel results for EQ_window_5

Table 8 presents the FE and RE regression results for EQ_window_5, providing insights into both within-bank and between-bank variations.

The Within R^2 values for both FE and RE models are notably low, suggesting that the current set of explanatory variables, in their linear forms, has limited ability to capture the time-series variation in EQ_window_5 *within* individual banks. However, the highly significant poolability test ($p < 0.001$) unequivocally confirms the presence of substantial unobserved bank-specific effects, thus justifying the essential use of a panel data framework as discussed in Section 3.5. In the RE model, the robust F-statistic is highly significant (p = 0.0002), indicating that the model as a whole has statistical explanatory power. Nevertheless, individual coefficients for the independent variables do not reach conventional levels of statistical significance. This phenomenon is often observed in panel datasets with a short time dimension (T=7, for the 2018-2024 period) and limited within-bank variation, where the bank-specific fixed effects or random effects account for a substantial portion of the variance.

Table 8: Panel results for EQ_window_5

Variable	FE coefficient (p-value)	RE coefficient (p-value)
Const	-0.0082 (0.000)	-0.0044 (0.9632)
SIZE	-0.0189 (0.8809)	-0.0382 (0.7036)
FEMA	0.0159 (0.6030)	0.0311 (0.3638)
FIEXP	0.1171 (0.5923)	0.0035 (0.8943)
NON	0.0014 (0.9859)	0.0021 (0.9599)
INDD	0.0177 (0.9029)	0.0125 (0.8525)
DFL	-0.1022 (0.1751)	-0.0175 (0.8494)
GDP	-0.0141 (0.5213)	-0.0095 (0.6097)
INF	0.0154 (0.6504)	0.0166 (0.4953)
BANK_SIZE_log_assets	-0.1564 (0.6181)	-0.2298 (0.1619)
Model Statistics	FE	RE
Within R^2	0.0163	-0.0056
Between R^2	0.1213	0.3880
Overall R^2	0.0873	0.2585
F-statistic p-value	0.9506	0.0956
Robust F-statistic p-value	1.0000	0.0002***
Poolability test p-value	0.0000***	—

***: p<0.001; p-values in parentheses

- Panel results for EQ_window_3

Table 9 reports the FE and RE regression results for EQ_window_3, offering a robustness check with a shorter-term earnings quality measure.

Table 9: Panel results for EQ_window_3

Variable	FE coefficient (p-value)	RE coefficient (p-value)
Const	-0.0079 (0.0000)	-0.0115 (0.8735)
SIZE	0.0382 (0.6780)	-0.0235 (0.7340)
FEMA	-0.0167 (0.4372)	0.0579 (0.2517)
FIEXP	0.5240 (0.0150)*	-0.0438 (0.0077)**
NON	0.0242 (0.5807)	-0.0367 (0.2906)
DFL	0.0880 (0.5061)	0.0572 (0.6088)
GDP	-0.0014 (0.9703)	-0.0062 (0.8562)
INF	-0.0067 (0.9171)	0.0090 (0.8883)
BANK_SIZE_log_assets	0.0246 (0.8920)	-0.1671 (0.0230)*

*: p<0.05; **: p<0.01; p-values in parentheses

The results for EQ_window_3 show that financial expertise (FIEXP) is statistically significant in both the FE (p = 0.015) and RE (p = 0.0077) models, although with opposite signs. This difference suggests that the effect of financial expertise may vary depending on whether the variation occurs within banks over time or across banks. Prior literature indicates that financial expertise enhances board monitoring and improves financial reporting quality (Klein, 2002; Xie et al., 2003).

Bank Size (BANK_SIZE_log_assets) remains negatively significant (p < 0.05) in the RE model, consistent with earlier findings and suggesting that larger banks may experience lower earnings quality due to greater operational complexity and reduced transparency (Alessandrini et al., 2010).

- Hausman test

To determine the appropriate panel model specification for EQ_window_5, a Hausman test was conducted, as described in Section 3.4. The test evaluates whether individual effects are correlated with the regressors, which would invalidate the Random Effects (RE) estimator. Table 10 reports the Hausman test yields a p-value of 0.3601, exceeding the conventional significance level of 0.05. Therefore, the null hypothesis cannot be rejected, indicating that the RE estimator is consistent. Based on this result, the Random Effects model is preferred due to its greater efficiency and its ability to capture both within-bank and between-bank variation in earnings quality.

Table 10: Hausman test results for EQ_window_5

Statistic	Value
Chi-square	8.7831
p-value	0.3601
df	8

4.5. Panel regression results and robustness using Driscoll–Kraay standard errors

To ensure robust inference, the panel models are re-estimated using Driscoll–Kraay standard errors. This method accounts for heteroskedasticity, serial correlation, and cross-sectional dependence, which are common in banking panel data.

Table 11 reports the results using EQ_window_5 as the main earnings quality measure. Bank size (BANK_SIZE_log_assets) has a negative and statistically significant effect in both fixed effects and random effects models. This confirms that larger banks tend to have lower earnings quality when measured over longer periods.

Other governance variables, including board size (SIZE), gender diversity (FEMA), and financial expertise (FIEXP), are not statistically significant in the linear panel models. GDP growth and inflation are also not significant.

As a robustness check, the models are re-estimated using EQ_window_3 as an alternative earnings quality measure. The Hausman test indicates that the random effects model is appropriate. Bank size remains statistically significant, but the coefficient becomes positive. This difference reflects the shorter time horizon captured by EQ_window_3 compared to EQ_window_5. This result indicates that the effect of bank size on earnings quality depends on the measurement horizon. When earnings quality is measured over a shorter window (EQ_window_3), larger banks may benefit from diversification, stable funding, and operational scale, which contribute to more stable earnings. However, when measured over a longer window (EQ_window_5), increased operational complexity and governance challenges may reduce earnings persistence and transparency. This finding suggests that bank size has both stabilizing and complexity-related effects, and its overall impact depends on the time horizon considered.

Table 11: Panel regression results using Driscoll–Kraay standard errors

Variable	FE (EQ_window_5)	RE (EQ_window_5)	FE (EQ_window_3)	RE (EQ_window_3)
SIZE	-0.0189	-0.0382	-0.0047	-0.0191
	(0.0439)	(0.0369)	(0.0394)	(0.0258)
FEMA	0.0159	0.0311	-0.0102	-0.0100
	(0.0284)	(0.0342)	(0.0163)	(0.0232)
FIEXP	0.1171	0.0035	-0.2187	-0.0381***
	(0.1859)	(0.0302)	(0.4038)	(0.0076)
DFL	-0.1022	-0.0175	0.0031	0.0019
	(0.0622)	(0.0976)	(0.0366)	(0.0373)
BANK_SIZE_log_assets	-0.1564***	-0.2298***	0.1249***	0.1258***
	(0.0454)	(0.0743)	(0.0303)	(0.0298)
GDP	-0.0141	-0.0095	-0.0051	-0.0031
	(0.0105)	(0.0250)	(0.0055)	(0.0046)
INF	0.0154	0.0166	-0.0009	-0.0008
	(0.0155)	(0.0238)	(0.0060)	(0.0058)
Observations	159	159	159	159
R^2 (Overall)	0.0873	0.2585	-0.4001	0.0461

***: p < 0.01

Financial expertise (FIEXP) is statistically significant in the random effects model for EQ_window_3, with a negative coefficient. Other governance variables remain insignificant. The robustness analysis confirms that bank size is an important determinant of earnings quality. The results also support the GAM findings, which show that governance variables have nonlinear effects.

4.6. Generalized additive models and nonlinear evidence

To examine nonlinear relationships between governance characteristics and earnings quality, this study estimates a Generalized Additive Model (GAM), as described in Section 3.5. This approach allows the effect of each variable to be determined by the data, without imposing a linear structure. As a result, GAM can capture nonlinear patterns and interaction effects that linear models may not detect (Friedman et al., 2000).

The GAM achieves a Pseudo R^2 of 0.6042. This value is substantially higher than the R^2 of 0.375 obtained from the best OLS model for EQ_window_5. This result shows that nonlinear effects play an important role in explaining earnings quality. It also confirms that GAM provides a better fit than linear models. This finding highlights the importance of using flexible modeling approaches when analyzing governance and earnings quality relationships.

Table 12 presents the estimation results of the GAM model. The smooth terms for SIZE, FEMA, DFL, and BANK_SIZE_log_assets are statistically significant. This result indicates that these variables have nonlinear effects on earnings quality.

Table 12: GAM results for EQ_window_5

Smooth term	EDoF	Ref. Df	F-statistic	p-value
s(SIZE)	9.900	9.900	3.245	0.021*
s(FEMA)	5.900	5.900	2.871	0.043*
s(SIZE, FEMA)	6.400	6.400	3.162	0.034*
s(FIEXP)	4.600	4.600	1.196	0.289
s(DFL)	8.100	8.100	3.593	0.008**
s(GDP)	4.500	4.500	0.985	0.460
s(INF)	1.200	1.200	0.817	0.496
s(BANK_SIZE_log_assets)	2.700	2.700	4.887	0.006**
Model statistics				
Pseudo R-Squared	0.6042			
GCV Score	0.4111			

EDoF refers to the estimated degrees of freedom for the smooth term, indicating its complexity; Ref. Df is the reference degrees of freedom; F-statistic and p-value assess the significance of the smooth term; **: p<0.01; *: p<0.05

The smooth term for FEMA is statistically significant (p = 0.043). Fig. 1 shows that earnings quality improves as FEMA increases from low to moderate levels. At higher levels, the effect becomes weaker and more stable. This result suggests that gender diversity improves governance effectiveness, although its marginal effect decreases at higher levels.

Board size (SIZE) also has a statistically significant nonlinear effect (p = 0.021). Fig. 2 shows the estimated relationship between board size and earnings quality. The results indicate that earnings quality initially increases slightly as board size increases from low levels. However, as board size increases further, earnings quality declines. At higher levels, the relationship becomes flatter, indicating a weaker marginal effect. This pattern confirms that the relationship between board size and earnings quality is nonlinear.

Importantly, the effect of board size depends on female board representation. Fig. 3 also shows the marginal effect of board size at different levels of

FEMA. Separate curves are presented for low, mean, and high levels of FEMA. These curves differ in both shape and slope, indicating an interaction effect. When FEMA is low, larger boards are associated with lower earnings quality. When FEMA is higher, this negative effect becomes weaker or disappears. This finding suggests that gender diversity improves board effectiveness and reduces coordination and monitoring problems in larger boards.

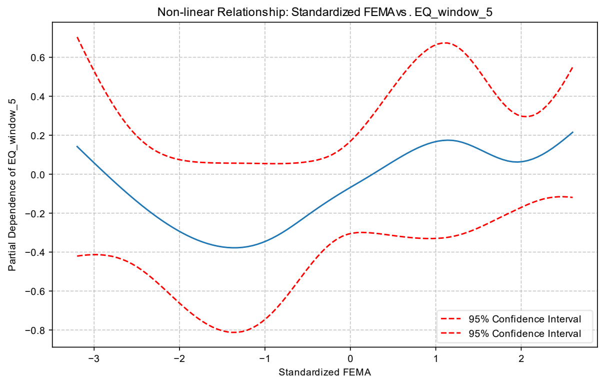

Fig. 1: Non-linear relationship: FEMA vs. EQ_window_5

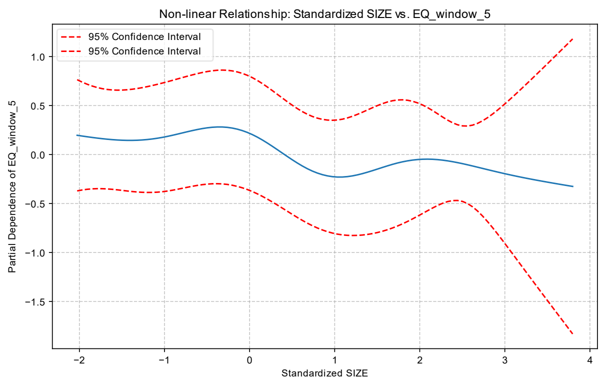

Fig. 2: Non-linear relationship: SIZE vs. EQ_window_5

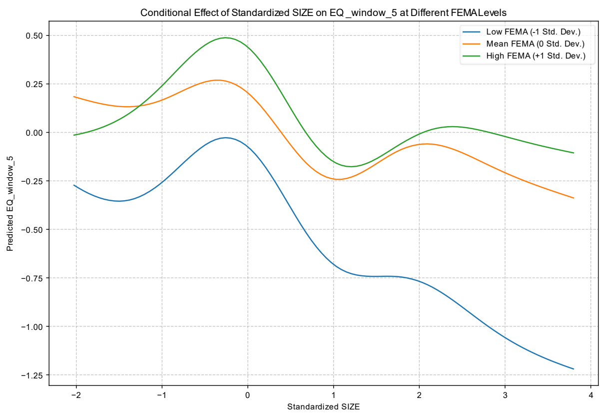

Fig. 3: Conditional effect of standardized SIZE on EQ_window_5 at different FEMA levels

Similarly, DFL exhibits a significant nonlinear relationship (p = 0.008). The partial dependence plot (Fig. 4) shows a positive association between DFL and earnings quality, with diminishing marginal effects at higher levels of deposit reliance. This result is consistent with prior findings that stable deposit funding enhances earnings stability and reporting quality (Hanson et al., 2015; Vazquez and Federico, 2015).

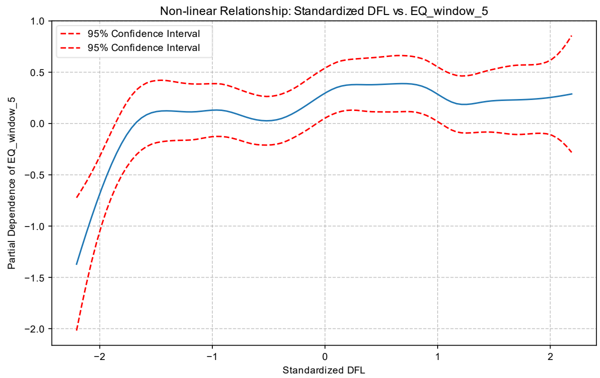

Fig. 4: Non-linear relationship: DFL vs. EQ_window_5

Bank size (BANK_SIZE_log_assets) has a statistically significant nonlinear effect (p = 0.006), as shown in Fig. 5. The results show that earnings quality declines at intermediate levels of bank size but stabilizes or improves for the largest banks. This pattern reflects both increased organizational complexity and stronger governance capacity in larger banks (Alessandrini et al., 2010).

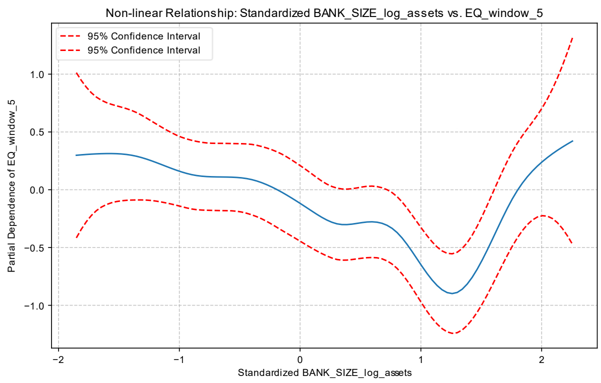

Fig. 5: Non-linear relationship: Bank_size_log_assets vs. EQ_window_5

In contrast, FIEXP, GDP, and INF are not statistically significant. Their nonlinear effects appear limited in this sample. In summary, the GAM results provide clear evidence of nonlinear and moderating governance effects. The higher Pseudo R^2 confirms that nonlinear models explain earnings quality better than linear models. The results also show that gender diversity moderates the effect of board size. These findings highlight the importance of considering nonlinear and interaction effects when studying earnings quality in Vietnamese commercial banks.

4.7. Discussion

This study examines the determinants of earnings quality in listed Vietnamese commercial banks during 2018–2024 using OLS, panel regression, and Generalized Additive Models. The results highlight the importance of bank characteristics and governance factors. They also show that nonlinear and interaction effects play an important role and cannot be fully captured by conventional linear models.

A consistent finding across OLS and panel models for EQ_window_5 is the negative association between bank size (BANK_SIZE_log_assets) and earnings quality. This result remains robust when using Driscoll–Kraay standard errors. This suggests that larger banks may face greater organizational complexity and information asymmetry, which can reduce reporting transparency (Alessandrini et al., 2010). However, the GAM results reveal a nonlinear relationship. Earnings quality declines at intermediate levels of bank size but stabilizes or improves among the largest banks. This finding suggests that while increasing size introduces governance and monitoring challenges, very large banks may benefit from stronger internal controls, diversified operations, and more effective oversight.

Deposit-based financial leverage (DFL) shows a positive and statistically significant effect in the OLS models for EQ_window_5, consistent with the view that stable deposit funding supports earnings stability and reporting reliability (Allen et al., 2015; Hanson et al., 2015). The GAM results confirm a nonlinear effect, suggesting that leverage improves earnings quality, but the marginal benefits decline at higher levels. This result indicates that deposit funding enhances governance and monitoring, but its effectiveness is not constant across all levels. The absence of significance for DFL in EQ_window_3 models suggests that these benefits are more evident over longer horizons (Vazquez and Federico, 2015).

Board gender diversity (FEMA) also shows a statistically significant nonlinear effect in the GAM model. In addition, the GAM results reveal that gender diversity moderates the effect of board size. The marginal effect analysis shows that larger boards are associated with lower earnings quality when female board representation is low. However, when female representation is higher, this negative effect becomes weaker or disappears. This finding suggests that gender diversity improves board effectiveness and reduces coordination and monitoring problems in larger boards. This result is consistent with prior research showing that gender diversity enhances governance and financial reporting quality (Srinidhi et al., 2011).

Board size (SIZE) itself also shows a significant nonlinear effect. The GAM results indicate that the effect of board size is not constant and depends on governance conditions. Larger boards may create coordination challenges, but their effectiveness improves when governance quality is stronger, particularly when female representation is higher. This finding highlights the importance of considering interaction effects when evaluating governance structures. Financial expertise (FIEXP) shows limited and inconsistent effects across model specifications. This suggests that financial expertise alone may not be sufficient to improve earnings quality. Its impact may depend on broader governance conditions and institutional context (Klein, 2002; Xie et al., 2003).

Macroeconomic variables, including GDP growth and inflation, are not statistically significant in the GAM model. This suggests that earnings quality is primarily influenced by bank-level governance and structural characteristics rather than macroeconomic conditions.

In summary, the GAM model provides important insights beyond linear models. The panel regression results remain consistent when using Driscoll–Kraay standard errors and alternative earnings quality measures. The GAM achieves higher explanatory power and captures nonlinear and moderating relationships. The results show that governance variables affect earnings quality in complex ways, and their effects depend on both scale and board composition. These findings confirm that flexible modeling approaches provide a more accurate understanding of earnings quality determinants (Friedman et al., 2000).

5. Conclusions

This study examines the determinants of earnings quality in Vietnamese commercial banks using linear, panel, and nonlinear models. The results show that earnings quality is mainly driven by bank-specific characteristics and governance factors. These findings remain robust when using Driscoll–Kraay standard errors and alternative earnings quality measures. In contrast, macroeconomic variables such as GDP growth and inflation do not have a significant effect during the sample period.

First, bank size has a nonlinear relationship with earnings quality. Linear models show a negative association, suggesting that larger banks face greater complexity and information asymmetry (Alessandrini et al., 2010). However, the GAM results show that this effect is not constant. Earnings quality declines at intermediate levels of bank size but stabilizes or improves among the largest banks. This finding suggests that while increasing size creates governance challenges, very large banks may benefit from stronger internal controls, better governance systems, and more effective regulatory oversight.

Second, deposit-based financial leverage has a positive effect on earnings quality. This result supports the role of stable deposit funding in improving financial reporting reliability (Allen et al., 2015; Hanson et al., 2015). However, the GAM results show that this effect is nonlinear. The marginal benefits decline at higher levels of leverage. This suggests that deposit funding improves monitoring and discipline, but its effectiveness becomes weaker beyond a certain level.

Third, board gender diversity has both nonlinear and moderating effects. The GAM results show that female board representation has a nonlinear relationship with earnings quality. In addition, gender diversity moderates the effect of board size. When female representation is low, larger boards are associated with lower earnings quality. When female representation is higher, this negative effect becomes weaker or disappears. This finding suggests that gender diversity improves board effectiveness and helps reduce coordination and monitoring problems in larger boards. This result is consistent with prior research showing that gender diversity improves governance and financial reporting quality (Srinidhi et al., 2011).

Methodologically, this study shows the importance of using nonlinear models and robust inference methods. The GAM model provides higher explanatory power than linear models and identifies nonlinear and interaction effects. The use of Driscoll–Kraay standard errors confirms that the main findings are robust and not driven by biased standard errors. These results show that flexible models provide a more accurate understanding of earnings quality determinants (Friedman et al., 2000).

These findings have practical implications for managers, regulators, and investors. Bank managers should pay attention to governance structure and board composition when improving reporting quality. Increasing gender diversity may improve board effectiveness, especially in larger boards. Regulators should consider differences in bank size and governance structure when designing supervisory policies. Investors should also consider governance characteristics when evaluating financial reporting quality.

5.1. Limitations and future research

This study has several limitations that provide directions for future research. First, the sample covers the period 2018–2024. Future studies may extend the sample period to examine long-term trends and structural changes.

Second, although this study includes key governance variables, other factors such as ownership structure, audit committee characteristics, and risk management roles may also affect earnings quality. Including these variables may provide a more complete understanding of governance effects.

Third, this study focuses on Vietnamese commercial banks. Future research may examine other countries to assess whether similar patterns exist in different institutional settings.

Finally, while GAM identifies nonlinear and interaction relationships, future studies may use alternative methods to examine causal mechanisms and institutional factors in more detail.

List of abbreviations

Adj. R^2	Adjusted coefficient of determination
BANK_SIZE_log_assets	Natural logarithm of total bank assets
DFL	Deposit-based financial leverage
DFL×BANK_SIZE_log_assets	Interaction term between deposit-based financial leverage and bank size
DUAL	Chief executive officer duality
EDoF	Estimated degrees of freedom
EQ	Earnings quality
EQ_window_3	Earnings quality index based on a three-year rolling window
EQ_window_5	Earnings quality index based on a five-year rolling window
FE	Fixed effects
FEMA	Female board ratio
FIEXP	Financial expertise ratio
GAM	Generalized Additive Model
GAMs	Generalized Additive Models
GCV	Generalized cross-validation score
GDP	Annual gross domestic product growth rate
HC1	Heteroskedasticity-consistent standard error estimator, type 1
INDD	Independent director ratio
INF	Annual inflation rate
N	Number of observations
NON	Non-executive director ratio
OLS	Ordinary Least Squares
p-value	Probability value
RE	Random effects
Ref. Df	Reference degrees of freedom
R^2	Coefficient of determination
SBV	State Bank of Vietnam
SD	Standard deviation
SIZE	Board size
SIZE_cube	Cubic board size term
SIZE_sq_std	Standardized squared board size term
SIZE_x_FEMA_std	Interaction term between standardized board size and female board ratio
VIF	Variance Inflation Factor
Z_PERS	Standardized earnings persistence score
Z_PRED	Standardized earnings predictability score
Z_SMOOTH	Standardized earnings smoothness score
Z_VOL	Standardized earnings volatility score

Acknowledgment

This work was supported by Ho Chi Minh University of Banking (HUB), Vietnam.

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Alessandrini P, Presbitero AF, and Zazzaro A (2010). Bank size or distance: What hampers innovation adoption by SMEs? Journal of Economic Geography, 10(6): 845-881. https://doi.org/10.1093/jeg/lbp055

Allen F, Carletti E, and Marquez R (2015). Deposits and bank capital structure. Journal of Financial Economics, 118(3): 601-619. https://doi.org/10.1016/j.jfineco.2014.11.003

Anagnostopoulou SC and Tsekrekos AE (2017). The effect of financial leverage on real and accrual-based earnings management. Accounting and Business Research, 47(2): 191-236. https://doi.org/10.1080/00014788.2016.1204217

Awuye IS and Aubert F (2022). The impact of leverage on earnings management and the trade-off between discretionary accruals and real earnings management. Journal of Accounting and Taxation, 14(1): 89-101.

Barth ME, Clinch G, and Israeli D (2016). What do accruals tell us about future cash flows? Review of Accounting Studies, 21: 768-807. https://doi.org/10.1007/s11142-016-9360-4

Bertay AC, Demirgüç-Kunt A, and Huizinga H (2013). Do we need big banks? Evidence on performance, strategy and market discipline. Journal of Financial Intermediation, 22(4): 532-558. https://doi.org/10.1016/j.jfi.2013.02.002

Bushman RM and Williams CD (2012). Accounting discretion, loan loss provisioning, and discipline of banks' risk-taking. Journal of Accounting and Economics, 54(1): 1-18. https://doi.org/10.1016/j.jacceco.2012.04.002

Bushman RM and Williams CD (2015). Delayed expected loss recognition and the risk profile of banks. Journal of Accounting Research, 53(3): 511-553. https://doi.org/10.1111/1475-679X.12079

Chang X, Huang J, Wang T, and Wang Y (2025). Board gender diversity and strategic change: Integrating information and social categorization perspectives. Humanities and Social Sciences Communications, 12: 1671. https://doi.org/10.1057/s41599-025-05930-5

Cheng Q, Warfield T, and Ye M (2011). Equity incentives and earnings management: Evidence from the banking industry. Journal of Accounting, Auditing & Finance, 26(2): 317-349. https://doi.org/10.1177/0148558X11401219

Coles JL, Daniel ND, and Naveen L (2008). Boards: Does one size fit all? Journal of Financial Economics, 87(2): 329-356. https://doi.org/10.1016/j.jfineco.2006.08.008

Dang C, Li ZF, and Yang C (2018). Measuring firm size in empirical corporate finance. Journal of Banking & Finance, 86: 159-176. https://doi.org/10.1016/j.jbankfin.2017.09.006

de Comères QB (2025). Predicting European banks distress events: Do financial information producers matter? International Review of Financial Analysis, 105: 104417. https://doi.org/10.1016/j.irfa.2025.104417

Dechow P, Ge W, and Schrand C (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. Journal of Accounting and Economics, 50(2-3): 344-401. https://doi.org/10.1016/j.jacceco.2010.09.001

Dechow PM and Dichev ID (2002). The quality of accruals and earnings: The role of accrual estimation errors. The Accounting Review, 77(s-1): 35-59. https://doi.org/10.2308/accr.2002.77.s-1.35

Di Martino G, Dicuonzo G, Galeone G, and Dell'Atti V (2017). Does the new European banking regulation discourage earnings management? International Business Research, 10(10): 45-56. https://doi.org/10.5539/ibr.v10n10p45

Dong Y, Teoh SH and Zhang Y (2025). Bringing operational context into earnings quality models: A new framework. https://doi.org/10.2139/ssrn.4664183

Friedman J, Hastie T and Tibshirani R (2000). Additive logistic regression: A statistical view of boosting (with discussion and a rejoinder by the authors). The Annals of Statistics, 28(2): 337-407. https://doi.org/10.1214/aos/1016218223

Hanson SG, Shleifer A, Stein JC, and Vishny RW (2015). Banks as patient fixed-income investors. Journal of Financial Economics, 117(3): 449-469. https://doi.org/10.1016/j.jfineco.2015.06.015

Hung DN and Van VTT (2020). Studying the impacts of earnings quality on stock return: Experiments in Vietnam. International Journal of Advanced and Applied Sciences, 7(4): 45-53. https://doi.org/10.21833/ijaas.2020.04.007

Klein A (2002). Audit committee, board of director characteristics, and earnings management. Journal of Accounting and Economics, 33(3): 375-400. https://doi.org/10.1016/S0165-4101(02)00059-9

Leuz C, Nanda D, and Wysocki PD (2003). Earnings management and investor protection: An international comparison. Journal of Financial Economics, 69(3): 505-527. https://doi.org/10.1016/S0304-405X(03)00121-1

Pham TP, Nguyen DN, Huynh TH, Le T, and Nguyen TX (2023). The effect of bank fintech on bank efficiency: An empirical study from Vietnam. Journal of Southwest Jiaotong University, 58(4): 447-472. https://doi.org/10.35741/issn.0258-2724.58.4.36

Saeed MS (2014). Bank-related, industry-related and macroeconomic factors affecting bank profitability: A case of the United Kingdom. Research Journal of Finance and Accounting, 5(2): 42-50.

Srinidhi BIN, Gul FA, and Tsui J (2011). Female directors and earnings quality. Contemporary Accounting Research, 28(5): 1610-1644. https://doi.org/10.1111/j.1911-3846.2011.01071.x

Tan C and Taufiik M (2022). Board diversity and financial reporting quality: Does firm size matter? Journal of Contemporary Accounting, 4(2): 80–94. https://doi.org/10.20885/jca.vol4.iss2.art2

Thi Tran TN (2024). Liquidity creation and banking stability: An approach using the Bayes method. Cogent Economics & Finance, 12(1): 2417760. https://doi.org/10.1080/23322039.2024.2417760

Vafeas N and Vlittis A (2024). Earnings quality and board meeting frequency. Review of Quantitative Finance and Accounting, 62: 1037-1067. https://doi.org/10.1007/s11156-023-01230-8

Vazquez F and Federico P (2015). Bank funding structures and risk: Evidence from the global financial crisis. Journal of Banking & Finance, 61: 1-14. https://doi.org/10.1016/j.jbankfin.2015.08.023

Xie B, Davidson III WN, and DaDalt PJ (2003). Earnings management and corporate governance: The role of the board and the audit committee. Journal of Corporate Finance, 9(3): 295-316. https://doi.org/10.1016/S0929-1199(02)00006-8