Proceedings of CONF-BPS 2026 Symposium: GenAI, Labour Markets, and the Economics of Human and Financial Capital
DOI: 10.54254/2754-1169/2026.LD31581

The Marginal Effects of Corporate Governance on Bank Financial Performance and Their Conditional Variations

Yanlin Zheng

School of Accounting, Guangdong University of Finance & Economics, Guangzhou, China
18127962759@163.com

Abstract. With the growing salience of Environmental, Social, and Corporate Governance (ESG), how banks can bolster financial performance through improved corporate governance has become a critical research agenda. This study investigates the marginal effect of corporate governance on banks' financial outcomes by centering on the nexus between ESG's governance (G) score and the return on total assets (ROA). It further explores the transmission mechanism through which governance investments shape bank performance and how these effects differ across bank types. Employing descriptive analysis, a panel regression model (two-way fixed effects), and a machine learning approach (random forest regression with five-fold cross-validation), the study analyzes data from multiple listed Chinese banks from 2009 to 2024. The findings reveal a significantly positive association between the G score and ROA, and a significantly negative association between the G score and the non-performing loan (NPL) ratio. An additional 10,000 yuan of governance investment per 10,000 yuan of assets leads to a 0.0002-percentage-point increase in ROA, with significant cross-institutional heterogeneity. State-owned banks exhibit stronger performance responses to governance improvements. The results underscore the imperative of constructing a system for evaluating the marginal benefits of governance investment and to monitoring its effectiveness dynamically in order to achieve more efficient resource allocation.

Keywords: corporate governance, financial risk, bank performance, Heterogeneity analysis

1. Introduction

Amid the global ESG paradigm shift, how banks can improve their performance through governance investment has become a pivotal concern at the intersection of regulatory oversight and strategic management. In recent years, a burgeoning body of scholarly work has explored the impact of ESG on corporate financial risk. For example, some adopt basic linear regression specifications to examine the impact of ESG on Tobin's Q, the KZ index, total factor productivity (TFP), and the Z-score to gauge financial risk [1]. Others apply standard panel models and panel quantile regression to analyze the effect of ESG on financial risk [2]. However, in China, only a few researchers have focused on the banking sector—a uniquely regulated and information-asymmetric industry—when studying how ESG influences financial risk. In contrast, a proliferating strand of international literature has emerged in this area. Bruno, Iacoviello, and Giannetti use a panel fractional response

DOI: 10.54254/2754-1169/2026.LD31581

model, a control-function approach, and a two-step QMLE with bootstrap standard errors to investigate the nexus between ESG and financial risk [3]. Cantero-Saiz, Polizzi, and Scannella conduct panel data regression to study how banks' asset quality is affected by ESG [4,5]. Existing literature predominantly focuses on overall ESG performance or general firms, and devotes insufficient attention to the single governance (G) dimension in the banking sector and its heterogeneous effects across distinct ownership configurations. It is also uncommon to integrate machine learning methodologies for predictive analysis and result validation. This research focuses on the marginal effect of corporate governance on banks' financial performance, examining the relationship between the governance (G) score in ESG and return on total assets (ROA), and exploring how governance investment influences bank profitability.

The panel regression analysis utilizes a two-way fixed effects model. By applying mean-centering, the model effectively mitigates the influence of confounding factors arising from time or individual variations. In the machine learning component, random forest regression and five-fold cross-validation are employed as complementary analytical tools. These two methods help capture the complex (including nonlinear) relationships between the independent and dependent variables, and enhance the reliability of the results through repeated training and an assessment of model generalizability. The findings contribute to helping regulators develop a differentiated and dynamic framework for evaluating governance investment, thereby supporting the dual objectives of financial stability and shareholder value maximization.

2. Methodology

2.1. Data sources and sample selection

This study's dataset is primarily derived from 43 domestic listed banks, spanning the time horizon from 2009 to 2024. The governance (G) scores and ratings are retrieved from the Huazheng (China Securities) database, while the financial data and indicators are sourced from the CSMAR database [6,7]. The initial sample comprises 432,418 observations and 65 financial and governance variables. First, the original variables were renamed and consolidated. Second, to reduce the potential influence of extreme values on the regression results, outlier adjustment was performed via winsorization of all continuous variables at the 1% and 99% quantiles. Missing values were addressed through listwise deletion. After data filtering and preprocessing, the final valid dataset for regression analysis consists of 381 bank-year observations.

2.2. Theoretical framework and research hypotheses

Agency theory posits that the separation of ownership and control gives rise to principal-agent conflicts between shareholders and managers, and such conflicts are further exacerbated in banks owing to their inherent high leverage and information opacity [8]. According to the resource-based view, imperfectly imitable governance resources can yield sustainable competitive advantages [9]. Based on this, the study proposes the following hypotheses:

H1: Governance investment is significantly and positively related to bank performance.

H2: This positive effect is stronger in private banks than in state-owned banks.

H3: Governance investment and bank performance exhibit an inverted U-shaped relationship, indicating the existence of an optimal level.

DOI: 10.54254/2754-1169/2026.LD31581

2.3. Variable definition and measurement

In the ESG framework, the "G" component refers to corporate governance, whose core evaluation dimensions encompass board structure, information disclosure and transparency, executive compensation, and compliance issues. ROA, or return on assets, gauges a bank's profitability as the ratio of net profit to total assets. NPL, the non-performing loan ratio, evaluates a bank's credit risk exposure, defined as the balance of non-performing loans divided by the total amount of loans.

Following existing studies on bank performance, this research incorporates a set of control variables capturing financial and ownership attributes. The financial variables include bank size (Size), leverage ratio (Lev), capital adequacy ratio (CAR), asset growth rate (Growth), loan-to-deposit ratio (LDR), and cost-to-income ratio (CIR). The detailed definitions and calculation methods of all variables are provided in Table 1.

Table 1. Variable definitions

Variable Name	Symbol	Calculation Formula	Definition
Bank Size	Size	Natural logarithm of total assets (lnTA)	Larger banks usually have stronger risk-diversification ability and potentially higher profitability.
Leverage Ratio	Lev	Total liabilities ÷ Total assets	Higher leverage indicates greater financial risk.
Capital Adequacy Ratio	CAR	(Core capital + supplementary capital) ÷ risk-weighted assets	A higher capital level reflects stronger financial stability.
Asset Growth Rate	Growth	(Total assets in the current period − total assets in the previous period) ÷ previous period total assets	Banks with higher growth may achieve higher returns but may also face higher risks.
Loan-to-Deposit Ratio	LDR	Total loans ÷ total deposits	Indicating a bank's liquidity and operational efficiency;
State Ownership	SOE	=1 if the bank is state-owned; otherwise =0	Governance efficiency may differ across ownership types.
Cost-to-Income Ratio	CIR	Operating expenses ÷ operating income	An indicator of cost management efficiency; lower values are preferred.

3. Model construction

3.1. Panel regression model

To account for unobserved individual heterogeneity and time-varying trends, this study utilizes a two-way fixed-effects (TWFE) panel regression model for estimation. The model specification draws on Bai's extension of the conventional two-way fixed effects framework, and its basic form is as follows [10]:

$$ROA_{it} = \alpha_0 + \beta_1 G_{it} + \gamma Controls_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

$$NPL_{it} = \alpha_0 + \beta_1 G_{it} + \gamma Controls_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (2)$$

ROA_{it} represents the return on assets of bank i in year t; NPL_{it} denotes the non-performing loan ratio; G_{it} refers to the corporate governance score; $Controls_{it}$ is the set of control variables,

DOI: 10.54254/2754-1169/2026.LD31581

including Size, Lev, CAR, Growth, LDR, CIR, and SOE; μ_i represents the individual fixed effects; λ_t denotes the time fixed effects; and ε_{it} is the random error term.

3.2. Heterogeneity analysis model

Adopting the methodological framework proposed by Aiken and West, interaction terms are incorporated to examine the moderating effects of the relevant variables [11].

$$ROA_{it} = \alpha_0 + \beta_1 G_{it} + \beta_2 (G_{it} \times SOE_{it}) + \gamma Controls_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (3)$$

ROA_{it} represents the return on assets of individual i in period t; G_{it} denotes the corporate governance score; SOE_{it} indicates whether the bank is state-owned; $G_{it} \times SOE_{it}$ refers to the interaction term used to test heterogeneity; and the meanings of $Controls_{it}$, μ_i, λ_t, ε_{it} remain the same as defined above.

3.3. Endogenous treatment method

To alleviate potential endogeneity arising from reverse causality between governance investment and bank performance, this study follows the approach of Angrist and Pischke by employing the one-period lag of governance intensity as the key explanatory variable for the regression analysis [12]. The model is specified as follows:

$$ROA_{it} = \alpha_0 + \beta_1 L.\, G_{it} + \gamma Controls_{it} + \mu_i + \lambda_t + \varepsilon_{it} \quad (4)$$

$L.\, G_{it}$ represents the one-period lag of the corporate governance score, and the meanings of all other symbols remain the same as previously defined.

3.4. Machine learning model

To further quantify the relative importance of each variable on bank performance, this study integrates a random forest regression model to perform feature importance analysis. This approach follows the recommendation of Mullainathan and Spiess [13].

Assume the sample dataset is denoted as $\{(X_i, Y_i)\}i = 1^n$,

$$X_i = (G_{it}, Size_{it}, Lev_{it}, CAR_{it}, Growth_{it}, LDR_{it}, CIR_{it}), Y_i = ROA_{it}或NPL_{it} \quad (5)$$

Random Forest Regression Model:

$$\hat{f}RF\left(X\right) = \frac{1}{B} \sum b = 1^B T\left(X; \Theta_b\right) \quad (6)$$

$\hat{f}RF\left(X\right)$ denotes the predicted value of the dependent variable generated by the random forest; B represents the number of trees; $T(X; \Theta_b)$ refers to the prediction made by the b-th decision tree; Θ_b and denotes the parameters of the b-th tree.

DOI: 10.54254/2754-1169/2026.LD31581

3.5. Nonlinear analysis

To investigate the potential nonlinear association between governance intensity and performance, this study follows Wooldridge's discussion on functional form and incorporates a quadratic term into the baseline linear specification. The specific model is specified as follows:

$$ROA = \beta_0 + \beta_1 X + \beta_2 X^2 + \beta_3 Controls + \varepsilon \quad (7)$$

Here, X^2 represents the squared term of governance intensity. If the coefficient β_2 is statistically significant and different from zero, it indicates a nonlinear relationship between X and Y. Specifically, when $\beta_1>0$ and $\beta_2<0$, the relationship takes an inverted U-shape; when $\beta_1>0$ and $\beta_2<0$, the relationship follows a U-shape [13].

4. Empirical analysis

4.1. Descriptive statistics

Table 2 presents descriptive statistics for 423 bank-year observations: the mean ROA is 0.915% (s.d. 0.225%), the NPL is 1.23% (s.d. 0.39%), and the G score is 85.99 (s.d. 5.56), suggesting substantial cross-sectional heterogeneity in governance investments.

Table 2. Descriptive statistics of bank indicators

	ROA	NPL	G Score	GIntensity	LnSize	Lev	SOE	CIR	CAR	Growth
Observations	423	394	423	417	423	423	423	405	394	423
Mean	0.00915	1.23264	85.986	5856.08390	3.34345	0.92667	0.32151	30.64064	13.56766	0.08969
Std. Dev.	0.00225	0.39212	5.5625	392.61337	0.05553	0.01185	0.46761	5.83434	2.11108	0.13404

4.2. Panel regression analysis

To test Hypothesis 1, a two-way fixed-effects (TWFE) panel regression model is estimated.

Table 3. Regression results of the effect of G-score on ROA

Dep. Variable:	ROA_demeaned		R-squared		0.378	
Model:	OLS		Adj. R-squared		0.365	
Method:	Least Squares		F-statistic		28.27	
Date:	Sat,20 Sep 2025		Prob(F-statistic)		2.62E-34	
Time:	11:21:43		Log-Likelihood		1868.2	
No. Observations	381		AIC:		-3718.5	
Df Residuals:	372		BIC:		-3683.0	
Df Model	8					
Variable	coef	std err	t	P>|t|	[0.025	0.975]
const	-0.018231	0.011422	-1.596	0.111	-0.040690	0.004228
G score	0.000065	0.000019	3.369	0.001	0.000027	0.000103
LnSize	0.010823	0.001921	5.635	0.000	0.007047	0.014600

DOI: 10.54254/2754-1169/2026.LD31581

Table 3. (continued)

Lev	-0.013309	0.011244	-1.184	0.237	-0.035419	0.008800
SOE	0.000237	0.000205	1.155	0.249	-0.000166	0.000639
NPL	-0.002120	0.000290	-7.320	0.000	-0.002690	-0.001551
CIR	-0.000013	0.000022	-0.587	0.557	-0.000056	0.000030
CAR	0.000033	0.000056	0.588	0.557	-0.000077	0.000143
Growth	0.005212	0.000813	6.408	0.000	0.003613	0.006812
Omnibus		6.586		Durbin-Watson		0.57
Prob(Omnibus)		0.037		Jarque-Bera (JB)		4.142
Skew		0.052		Prob(JB)		0.12603
Kurtosis		2.500		Cond. No.		15297

Based on the regression results, the coefficient of the G-score on ROA is 0.000065 (p-value = 0.001), which is statistically significant at the 1% level, indicating a significantly positive relationship between corporate governance quality and bank performance. Among the control variables, LnSize and Growth exert significant positive effects on ROA, while NPL exhibits a significant negative association. The model's R^2 is 0.378, indicating that the set of explanatory variables explains 37.8% of the total variation in ROA. Therefore, Hypothesis 1 is supported.

Table 4. Regression results of the impact of governance intensity on ROA

Dep. Variable:	ROA_demeaned		R-squared		0.371	
Model:	OLS		Adj. R-squared		0.362	
Method:	Least Squares		F-statistic		44.16	
Date:	Sat,20 Sep 2025		Prob(F-statistic)		8.88E-36	
Time:	11:21:43		Log-Likelihood		2006.4	
No. Observations	381		AIC:		-4001	
Df Residuals	375		BIC:		-3977	
Df Residuals:	375					
Df Model	5					
Variable	coef	std err	t	P>|t|	[0.025	0.975]
const	8.1323E-20	0.0000645	1.26E-15	1	0	0
G score	6.527635E-05	1.937718E-05	3.369	0.001	2.717381E-05	1.033789E-04
GIntensity_demeaned	0.000001602	0.000000499	3.208	0.001	0.00000062	0.00000258
LnSize_demeaned	-0.0375	0.006	-6.271	0	-0.05	-0.025
Lev_demeaned	0.0079	0.009	0.771	0.441	-0.012	0.028
CIR_demeaned	-0.00009971	0.0000239	-4.165	0	0	-0.0000526
CAR_demeaned	-0.0002	0.0000491	-3.228	0.001	0	-0.0000619
Omnibus		7.158		Durbin-Watson		0.55
Prob(Omnibus)		0.028		Jarque-Bera (JB)		11.055
Skew		0.041		Prob(JB)		0.00398
Kurtosis		3.83		Cond. No.		28800

DOI: 10.54254/2754-1169/2026.LD31581

Notes: [1]Standard Errors assume that the covariance matrix of the errors is correctly specified [2].. The condition number is large, 2.88e+04. This might indicate that there is strong multicollinearity or other numerical problems.

Based on the regression results in Table 3, the coefficient of governance intensity is 0.000002 with a p-value of 0.0014, which is statistically significant, indicating that governance investment generates marginal returns. Specifically, for every additional 10,000 yuan of governance expenditure per 100 million yuan of total assets, ROA rises by 0.000002 percentage points. Given the relatively large condition number, potential issues of severe multicollinearity may exist. However, since the overall model is significant and the core explanatory variable remains statistically robust, the core conclusions remain robust.

4.3. Heterogeneity analysis results

Heterogeneity analysis results reveal that substantial disparities exist in the marginal returns of governance investments. For private banks, the coefficient of governance investment is 0.000001581 with a p-value of 0.014, achieving statistical significance at the 5% level; whereas for state-owned banks, the coefficient is 0.000000866 with a p-value of 0.293, which is statistically insignificant. In terms of economic significance, for private banks, an additional 1,000 yuan of governance investment per 10,000 yuan of assets increases ROA by approximately 0.0016 percentage points, representing approximately 1.8 times the magnitude of the effect observed for SOBs (0.0009 percentage points). Therefore, Hypothesis 2 is supported.

4.4. Stability and endogeneity tests

To ensure the credibility of the core findings, this study performs robustness checks and addresses endogeneity concerns. First, following the exclusion of outliers from the subsample, the positive effect of governance intensity remains statistically significant at the 5% level (coefficient = 0.000002107, p = 0.012); second, after converting governance intensity into logarithmic form, its coefficient (0.005542) exhibits no qualitative changes in sign or statistical significance, confirming that the results are not sensitive to the functional form of the variable. To alleviate the potential confounding effects of endogeneity on the results, the one-period lag of governance intensity (one period) is used for regression analysis. Its coefficient (0.000001261) remains significant at the 1% level (p = 0.011). Although the absolute value decreases, it offers robust empirical support for the causal claim that "governance investment drives bank performance improvement". The series of tests jointly indicate that the core conclusion of this study—namely, that governance investment enhances bank performance—is robust and credible.

4.5. Machine learning analysis

Random forest feature importance analysis yields a model performance of mean squared error (MSE) = 0.000003 and R^2 = 0.3973. This goodness of fit suggests that the model's explanatory power is comparable to that of the linear regression model, and the explanatory variables have stable predictive power for bank performance. Ranking these features by importance yields: Lev: 0.2612, Size: 0.2388, CAR: 0.2040, CIR: 0.1174, GIntensity: 0.1104, SOE: 0.0682. This reflects that leverage, bank size, and capital adequacy ratio are the three most critical features for predicting bank performance, revealing the inherent high-risk characteristics of the banking industry as a financial sector. Meanwhile, the importance of the core variable of this study—governance intensity—is significantly greater than that of ownership type, indicating that governance investment is one of the

DOI: 10.54254/2754-1169/2026.LD31581

most important factors affecting bank performance, aside from fundamental financial indicators and ownership characteristics.

4.6. Nonlinear relationship analysis

To investigate the potential nonlinear association between governance quality and bank performance posited earlier, a quadratic regression model is further estimated. The analysis results show that the quadratic model produces a coefficient of –0.000025 for the linear governance term and a coefficient of 0.000000 for the quadratic term. In the log-transformed regression, the governance coefficient is 0.005542 and statistically significant ($p = 0.0013$), which is statistically significant, thereby confirming a significant inverted U-shaped relationship between governance investment and bank performance—i.e., diminishing marginal returns. According to model predictions, the optimal governance investment level that maximizes bank ROA falls within the range of 47-50 million yuan per 1 billion yuan of total assets. This finding supports Hypothesis 3, indicating the existence of an optimal level of governance investment.

5. Conclusion

The findings of this study demonstrate that the G-score is significantly and positively associated with ROA, and significantly and negatively associated with NPL. Furthermore, distinct heterogeneity exists: the impact of the G-score on performance is larger for state-owned banks. The fundamental differences between state-owned and private banks in terms of ownership, governance and resources result in markedly divergent marginal effects of governance quality on performance. Private banks, with their clear residual claims and market-based incentives, see governance investments directly translated into performance. State-owned banks, however, face multiple policy objectives and soft budget constraints, where governance resources are easily diluted by social responsibilities, compounded by lengthy decision-making chains and sluggish responsiveness. Regulatory approaches should be differentiated: for private banks, adopt a market-oriented framework permitting flexible remuneration and tax incentives; for state-owned banks, implement rule-based governance by strengthening board independence, imposing remuneration caps and deferred clawbacks, and balancing incentives with constraints to avoid inefficient one-size-fits-all solutions. However, the study still has limitations. The dataset is predominantly composed of listed banks, which may not fully represent the entire banking sector—particularly small and medium-sized non-listed banks. In addition, the G-score may not fully capture the comprehensive level of corporate governance. To address these limitations, future research could incorporate data from non-listed banks, explore the underlying mechanisms through which the G-score influences bank performance, and more deeply explore the complex nonlinear relationship between governance quality and bank performance.

References

[1] Wang, L., Lian, Y., & Dong, J. (2022). The Mechanism through Which ESG Performance Influences Corporate Value. Securities Market Herald, (05), 23–34.

[2] Xie, C., & Li, W. (2023). Can Firms' Improved ESG Performance Reduce Financial Risk? Empirical Evidence from Chinese Listed Companies. Journal of Hunan University (Social Sciences), 37(02), 51–58. https://doi.org/10.16339/j.cnki.hdxbskb.2023.02.008

[3] Cantero-Saiz M, Polizzi S, Scannella E. ESG and asset quality in the banking industry: The moderating role of financial performance [J]. Research in International Business and Finance, 2024, 69: 102221.

DOI: 10.54254/2754-1169/2026.LD31581

[4] Bruno E, Iacoviello G, Giannetti C. Bank credit loss and ESG performance [J]. Finance Research Letters, 2024, 59: 104719.
[5] datasource, 2009-2024, Huazheng (China Securities) database, https: //idmt.chindices.com
[6] datasource, 2009-2024, CSMAR database, https: //data.csmar.com/
[7] Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics, 3(4), 305–360.
[8] Barney, J. B. (1991). Firm resources and sustained competitive advantage. Journal of Management, 17(1), 99–120.
[9] Bai, J. (2009). Panel data models with interactive fixed effects. Econometrica, 77(4), 1229–1279.
[10] AIKEN L S, WEST S G. Multiple regression: Testing and interpreting interactions [M]. Sage, 1991.
[11] Angrist, J. D., & Pischke, J. S. (2009). Mostly harmless econometrics: An empiricist's companion. Princeton University Press.
[12] Mullainathan, S., & Spiess, J. (2017). Machine learning: An applied econometric approach. Journal of Economic Perspectives, 31(2), 87–106.
[13] WOOLDRIDGE J M. Introductory econometrics: A modern approach [M]. 6th ed. Boston: Cengage Learning, 2016.