

ESG, Financial Constraints, and Green Technology Innovation: Evidence from China with the Moderating Role of Environmental Public Attention

Jiawen Wang^{1*}

¹ Faculty of Finance, City University of Macau, Macao 999078, China

*Corresponding author: Jiawen Wang

Abstract

Against the backdrop of China's "dual-carbon" objectives and ongoing improvements in ESG disclosure standards, green technological innovation has become central to firms' sustainable development. Using a sample of Chinese A-share listed companies, this study examines how ESG performance affects firms' green technology innovation and whether the effect varies across regions and ownership types. To unpack the underlying channels, the analysis incorporates financing constraints as a mediating mechanism and environmental public attention as a moderating factor. The results show that ESG performance is positively associated with green innovation, with pronounced geographic variation and firm-level heterogeneity. The effect is stronger for non-state-owned enterprises than for state-owned enterprises. Financing constraints mediate the ESG–green innovation relationship, suggesting that improved ESG performance can ease funding frictions and thereby support innovation activities. In addition, environmental public attention significantly strengthens the positive effect of ESG on green innovation: in regions with higher public attention to environmental issues, the ESG-innovation linkage becomes more pronounced. Overall, this study contributes to the literature by integrating spatial heterogeneity and public environmental attention into the empirical analysis of ESG-driven green innovation.

Keywords

ESG, green technology innovation, financial constraint, environmental public attention

1. Introduction

Green technology innovation plays a critical role in facilitating the United Nations Sustainable Development Goals and the issue of climate change [1]. Although the world has made tremendous efforts to reduce the effects of climate change, many challenges still remain across the globe, and the situation in Asia seems even more severe [2]. China has put forth the idea for the establishment of an innovation system for market-oriented green technology with the goal of developing and enhancing clean energy, clean production, and energy-saving environmental protection industries [3]. China's efforts demonstrate its commitment to making green innovation integral to its development strategy and to pursuing collaborative as well as industrial innovation. The initiatives have resulted in leaps and bounds in strategic industries such as renewable energy and new energy vehicles.

During the implementation period of China's 14th Five-Year Plan, the nation's ESG development entered a phase of remarkable expansion and maturation. This transformative period witnessed substantial advancements across multiple dimensions such as the establishment of more sophisticated ESG governance frameworks, the accelerated development of comprehensive ESG standards, a notable surge in ESG-oriented investment activities, progressive improvements in corporate ESG management practices, and the continual refinement of ESG rating and assessment methodologies. This comprehensive evolution has fundamentally reshaped the corporate landscape, as Chinese enterprises have progressively deepened their understanding of ESG principles and begun integrating these considerations systematically into their operational strategies and strategic decision-making frameworks. Existing academic research has documented that superior ESG performance can facilitate corporate green innovation through two primary pathways: the signalling mechanism and the resource integration channel [3-6]. The signalling mechanism operates by conveying credible information about corporate sustainability commitments to external stakeholders, thereby reducing information asymmetry. Meanwhile, the resource integration channel functions by improving access to critical resources, including financial capital, human talent, technological capabilities, and collaborative partnerships that are essential for innovation activities.

Nevertheless, green innovation initiatives typically exhibit three distinctive characteristics that present substantial challenges for enterprises. They demand considerable financial investment upfront, higher levels of uncertainty and risk, and require extended development cycles before generating tangible returns. Consequently, corporations pursuing green innovation frequently encounter significant financing constraint bottlenecks that impede their innovation capacity. The capital-intensive nature of green technology development, combined with information asymmetries between firms and external investors, creates barriers to securing adequate funding for long-term sustainability projects. ESG performance can affect corporate green innovation, while results can be diverse and complex. This is mainly attributed to the difference in the nature of companies and unbalanced regional development. Thus, investigating the mechanism of ESG impact on green technology innovation is practically and theoretically meaningful. While most studies prove heterogeneous characteristics, discussions on regional gradient differences and their causes remain relatively broad. This study attempts to enrich current research on ESG and green innovation in emerging markets with a discussion on the role of financing constraints and regional differences and a closer look at the moderating effect of environmental public attention.

The contribution of this research mainly includes three aspects. First, it attempts to broaden the contextual applicability of studies on the relationship between ESG performance and green technology innovation. This provides supplementary empirical evidence based on the Chinese market to inform research in emerging market economies. In addition, it refines the dimensions of heterogeneity analysis, revealing regional gradients and enterprise nature differences, thereby enriching empirical cases in this field. Last, it provides empirical insights from the Chinese market for studies on the effect of financial constraints and environmental public attention.

2. Hypothesis Development

2.1 ESG and green technology innovation

ESG apparently aligns with the “dual carbon” goals. According to signalling theory, companies fulfilling ESG responsibilities can signal high-quality development to stakeholders, alleviate regulatory pressure, reduce public relations risks, and gain external support to improve green technological innovation [7]. Together with stakeholder theory, firms integrating ESG into their business activities can easily increase information transparency and therefore reduce information asymmetry with external stakeholders [8-9]. Overall, strong ESG performance reflects multidimensional positive development, strengthens stakeholder partnerships, sustains resources and promotes green technological innovation. Therefore, this study establishes the following hypothesis:

H1: As a firm's ESG performance increases, its green technology innovation increases.

2.2 The mediating effect of financial constraints

Good ESG performance conveys positive market signals. In other words, information asymmetry between firms and other stakeholders is reduced. The attraction of capital from investors, financial institutions and government policy incentives offsets financing constraints for the company [10]. Lowering the level of financing constraints allows the company to access sufficient financial capital to undertake green technology innovation activities, thus filling the funding gap and increasing the intensity of green technology innovation [11]. Conversely, companies with low ESG performance face high financing constraints because of amplified information asymmetry, thus hindering green technological innovation activities. Therefore, this study designs the following hypothesis:

H2: ESG performance reduces financial constraints, boosting business green technology innovation.

3. Data and Research Design

This study draws on a panel of Chinese A-share listed firms over the period 2010–2022. Firm-level ESG ratings are obtained mainly from the Sino-Security Index Information Service and iFinD, with complementary financial information collected from the CSMAR and Wind databases. To ensure data quality and improve the reliability of the estimates, this study implements the following procedures: (i) ST and *ST firms are removed; (ii) firms in the financial industry are excluded; (iii) observations with missing key variables are dropped; and (iv) all continuous variables are winsorized at the 1st and 99th percentiles to reduce the impact of outliers.

3.1 Variables

Green innovation (GI) is the dependent variable and sourced from CNRDS database. Following prior research, GI is measured as the natural logarithm of one plus the number of green invention patents. The same calculation method also applies to the alternative dependent variable in robustness checks, defined as the green utility model patent (GIn).

The primary independent variable is ESG performance, proxied by the Sino ESG Score. This composite rating evaluates firms' environmental, social, and governance practices based on publicly disclosed information. In robustness checks, this research employs the Bloomberg ESG Composite Score (ESG_B) as an alternative proxy.

Existing studies find that numerous factors influence corporate green innovation. To reduce omitted variable bias, this study selects the following core variables as controls: firm size (Size), asset structure (Lev), tangible assets (Tan), profitability (Roe), firm age (Age), shareholder ownership (Top1), nature of ownership (Soe) and growth capacity (Grow). Additionally, this study controls for year and firm fixed effects.

3.2 Research Design

To empirically test the hypothesis, the regression model is constructed as follows:

$$GI_{i,t} = \beta_0 + \beta_1 ESG_{i,t} + \theta Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \quad (1)$$

where $GI_{i,t}$ denotes the green innovation of firm i in year t and $ESG_{i,t}$ represents the firm's ESG performance. The term $Controls_{i,t}$ includes firm-level characteristic control variables. This research also includes firm effect μ_i and year λ_t fixed effects.

4. Results

4.1 Descriptive statistics

Table 1 reports the descriptive statistics for 11045 observations. The data reveal that the mean values for green innovation indicators (GI and GIn) are 0.305 and 0.241, respectively. Large standard deviations also indicate a significant uneven distribution and polarization in green innovation output across the sample firms. In contrast, the mean ESG score is 74.589 with a smaller standard deviation, suggesting that overall ESG performance is relatively stable and concentrated. Regarding control variables, state-owned enterprises

account for over half of the sample (53%), the average shareholding of the largest shareholder is 37.5%, and the mean leverage ratio is 47.3%. This reflects that the sample is primarily composed of mature enterprises with relatively concentrated ownership and moderate financial leverage, although significant individual variations remain in terms of profitability (Roe) and growth (Grow).

Table 1. Descriptive Statistics

	N	Mean	SD	Min	Max
GI	11045	0.305	0.705	0	2.996
GIn	11045	0.241	0.593	0	2.639
ESG	11045	74.589	5.383	56.93	84.1
Size	11045	22.499	1.440	17.893	25.596
Lev	11045	0.473	0.199	0.05	0.895
Roe	11045	0.047	0.060	-0.24	0.218
Age	11045	2.446	0.730	0	3.367
Grow	11045	0.184	0.423	-0.574	2.891
Top1	11045	0.375	0.161	0.087	0.752
Soe	11045	0.53	0.499	0	1
Tan	11045	0.227	0.176	0.002	0.687

4.2 Main results

The empirical results presented in Table 2 provide crucial insights into the relationship between corporate ESG and GI. Examining Model (1), which represents the baseline specification incorporating only the core ESG dimensional indicators without additional control variables, this research observes a regression coefficient of 0.00391 for ESG. This coefficient achieves statistical significance at the 1% level, indicating an exceptionally low probability that the observed relationship occurred by random chance. Specifically, the coefficient magnitude suggests that each unit increase in the comprehensive ESG score corresponds to approximately 0.00391 units of enhancement in green innovation metrics. This baseline result establishes a preliminary empirical foundation suggesting that corporate commitment to ESG principles serves as a meaningful driver of sustainable technological advancement. However, the interpretation of Model (1) requires careful consideration of its limitations. Omitting control variables can cause a serious omitted-variable bias. In that case, the link between ESG performance and innovation may not be a true ESG effect. Instead, it may be driven by other unobserved factors. To deepen the research conclusion, model (2) introduces a series of control variables to enhance the robustness of the analysis. In this extended model, the coefficient value of ESG is 0.00341, and its significance level reaches the statistical standard of 1%, indicating that after controlling for other potential influencing factors, ESG performance still has a significant positive effect on GI.

Table 2. Baseline Regression

	(1)	(2)
	GI	GI
ESG	0.00391*** (3.499)	0.00341*** (3.036)
Size		0.02478*** (2.608)
Lev		-0.02086 (-0.464)
Roe		0.12874 (1.452)
Age		0.03356 (1.569)
Grow		-0.00640 (-0.806)
Top1		-0.09564

		(-1.445)
Soe		-0.00847
		(-0.256)
Tan		-0.08844
		(-1.438)
cons	0.01383	-0.52282**
	(0.166)	(-2.338)
N	11045	11045
Firm FE	Y	Y
Year FE	Y	Y
r2	0.74338	0.74391

* All models are set to have firm fixed effects and year fixed effects. t statistics in parentheses* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. These apply to all regression results in this research.

4.3 Robustness Test

Two methods are adopted in the robustness test. The first method substitutes the core independent variable. The original ESG indicator is replaced with the Bloomberg ESG Composite Score (denoted as ESG_B), while the dependent variable remains the corporate green innovation (GI) indicator (corresponding to Model (1) in Table 3). The regression analysis results reveal a considerably positive coefficient at the 1% significance level, demonstrating the robustness of the baseline model's conclusion. Conversely, the second method substitutes the dependent variable. The GI is substituted by the logarithm of one plus the quantity of green utility model patents (GIn). The core explanatory variable remains the same. Referring to Model 2 in Table 3, ESG is significantly positive at the 5% significance level. This further proves that the core conclusion remains robust after changing the measurement method for GI. Using two alternative indicator substitutions, the robustness checks provide consistent results, confirming that ESG performance is significantly and positively associated with firms' green innovation (GI). This consistency indicates that the core conclusions are highly reliable.

Table 3. Robustness test

	(1)	(2)
	GI	GIn
ESG B	0.00607***	
	(4.925)	
Size	0.02252**	0.01370
	(2.383)	(1.469)
Lev	-0.01984	-0.00976
	(-0.443)	(-0.223)
Roe	0.11716	0.06204
	(1.319)	(0.726)
Age	0.03145	0.05296**
	(1.476)	(2.504)
Grow	-0.00763	-0.00854
	(-0.969)	(-1.122)
Top1	-0.09869	-0.26743***
	(-1.491)	(-3.949)
Soe	-0.00776	-0.00514
	(-0.234)	(-0.164)
Tan	-0.10268*	-0.10403*
	(-1.667)	(-1.822)
ESG		0.00237**
		(2.159)
cons	-0.38530*	-0.24348
	(-1.804)	(-1.092)
N	11045	11045
r2	0.74460	0.66356

5. Additional Discussion

The relationship between ESG performance and corporate green innovation demonstrates considerable complexity and variability in practical contexts. This heterogeneity stems primarily from two sources: fundamental differences in organizational characteristics (such as ownership structure, firm size, industry sector, and development stage) and pronounced disparities in regional economic development levels, institutional environments, and policy support mechanisms across China's vast geographic landscape. The eastern coastal regions, for instance, typically enjoy more developed financial markets, stronger regulatory enforcement, and greater stakeholder awareness compared to central and western regions, potentially leading to differential effects of ESG initiatives on innovation outcomes.

Given these complexities, investigating the mechanisms through which ESG performance influences green technology innovation holds substantial practical and theoretical significance. From a practical standpoint, understanding these mechanisms can guide corporate managers in formulating more effective ESG strategies and help policymakers design targeted interventions to promote sustainable innovation. Theoretically, such investigation enriches the comprehension of how nonfinancial performance dimensions interact with innovation processes in emerging market contexts characterized by institutional transitions and development imbalances.

5.1 Regional differences and enterprise nature

The heterogeneity analysis shown in Table 4 discusses the effect of ESG performance on GI in different groups of samples. To examine the heterogeneity characteristics of the regional differences, eastern, middle and western groups are formed to compare their distinct results. As shown in Table 4, Column (1), the estimated ESG coefficient for firms in the eastern region is 0.00277, which is significant at the 10% level. For the central regional group, the ESG coefficient in Table 4, Column (2) is 0.00237, and it is not significant. These findings can be attributed to the eastern region's advanced marketization and the dense concentration of green innovation resources. Therefore, it allows for the easy alignment of ESG practices with the market's needs for green innovation, resulting in the promotion of GI. Last, in Table 4, Column (3), for the western regional group, the ESG coefficient is 0.006 and has a positive statistical significance at the 5% level. This finding can be explained by the fact that in weaker information environments, ESG is more likely to be recognized as a credible signal for investors and financial institutions. Therefore, firms with better ESG in western area find it easier to gain financial support and facilitate more efficient resource allocation to green innovation activities.

As reported in Column (4), the SOE subsample provides an ESG coefficient of 0.00168, which becomes statistically insignificant. In contrast, in Column (5) for non-SOEs, the ESG coefficient is 0.00429 and remains significant (1% significance level). It is obvious that non-state-owned corporations are subject to stronger competitive pressure from the marketplace, and better ESG performance can be seen as a good signal of firms' reputation.

This evidence highlights an asymmetric effect of ESG, indicating that its ability to foster green innovation is contingent upon the nature of the firm's ownership. It varies across regions and firm types. There is a need for the development of differentiated strategies in future policy initiatives, for example, strengthening the synergetic support mechanisms for ESG and green innovation in central areas.

Table 4 Heterogeneity Analysis Results

	(1) East	(2) Middle	(3) West	(4) State-owned	(5) Non-State-owned
	GI	GI	GI	GI	GI
ESG	0.00277*	0.00237	0.00600**	0.00168	0.00429***
	(1.958)	(0.926)	(2.238)	(0.998)	(2.777)
Controls	Yes	Yes	Yes	Yes	Yes
cons	-0.25902	-0.81138	-0.63117	-0.08024	-1.05677***
	(-0.907)	(-1.497)	(-1.194)	(-0.226)	(-3.373)
N	7535	1939	1571	5859	5186
r2	0.76248	0.72699	0.62660	0.75661	0.73358

5.2 The mediator-financial constraint

This study employs the SA index (absa) as a quantitative measure to assess the severity of corporate financing constraints. The SA index exhibits a positive correlation with challenges of funding availability. Specifically, elevated index values signal more pronounced financing obstacles, reflecting heightened difficulties in accessing credit markets, unfavourable borrowing terms, or restricted cash flow positions that impede innovation investment capacity. The empirical findings documented in Table 5 provide several critical insights into this relationship. The first column of results confirms that ESG performance exerts a statistically significant and positive effect on GI. More importantly, the second column presents compelling evidence regarding the mediating mechanism. The results demonstrate that superior ESG performance substantially alleviates corporate financing constraints, as evidenced by a regression coefficient of -0.00135 that achieves statistical significance at the rigorous 1% level. This reduction translates directly into enhanced capital acquisition capability, improved access to diverse funding sources, and greater financial flexibility to support long-term green innovation projects that typically require patient capital and sustained resource commitment. From a theoretical perspective, corporate ESG commitment functions as a credible signal transmitted to capital markets. This signalling mechanism narrows the information gap between investors and firms, enabling external capital providers to more clearly identify the enterprise's long-term development value, which in turn reduces information friction costs during the financing process.

Specifically, green innovation is typically associated with high barriers, long-term innovation project durations, and high risk. From the resource-based theory, capital, being an important part of a core resource of a firm, suggests that easing funding constraints essentially optimizes and upgrades a firm's resource base. Sufficient capital is necessary for a firm to deliver an end-to-end innovation solution support. In this case, lower financing costs enable firms to enter more advanced and better-equipped areas of low-carbon factor inputs. Accordingly, financing constraints mediate the effect of ESG performance on GI.

Table 5. Financial Constraint (SA Index)

	(1)	(2)
	GI	absa
ESG	0.00341***	-0.00135***
	(3.036)	(-7.957)
Size	0.02478***	-0.02869***
	(2.608)	(-7.803)
Lev	-0.02086	-0.00976
	(-0.464)	(-1.026)
Roe	0.12874	0.04230**
	(1.452)	(2.191)
Age	0.03356	0.02861***
	(1.569)	(6.551)
Grow	-0.00640	0.00374
	(-0.806)	(1.624)
Top1	-0.09564	0.03811**
	(-1.445)	(2.492)
Soe	-0.00847	0.00486
	(-0.256)	(1.096)
Tan	-0.08844	0.00541
	(-1.438)	(0.422)
_cons	-0.52282**	4.46071***
	(-2.338)	(56.550)
N	11045	11045
r2	0.74391	0.96262

5.3 The moderator: environmental public attention

Apart from the internal organizational drivers that promote green innovation, external environmental factors can fundamentally shape and reconfigure the ESG-innovation relationship. This research primarily investigates the moderating effect of public environmental attention on the relationship between corporate ESG performance and green technology innovation outcomes. Public environmental awareness works as a crucial indicator that captures societal preferences regarding ecological sustainability, environmental protection, and responsible corporate behaviour. This collective consciousness reflects not only passive concern but also active engagement with environmental issues that can translate into tangible pressure on corporate actors. In the era of big data, online search behaviour records provide an unprecedented opportunity to capture and quantify the attention that market participants, including consumers, investors, media outlets, and regulatory bodies, direct towards specific environmental events, concerns, and corporate practices. These digital footprints offer real-time insights into evolving preferences, emerging priorities, and behavioural intentions across diverse stakeholder groups. The theoretical rationale underlying this measurement approach posits that search activity reflects genuine interest and concern that can subsequently influence decision-making processes, consumption patterns, and investment choices.

Cities or regions with higher levels of public attention to environmental concerns cultivate an institutional and cultural context wherein firms operating within these areas develop heightened awareness of sustainable development imperatives. This environmental consciousness enlarges reputational risks associated with poor environmental performance, strengthens stakeholder expectations for corporate ecological responsibility, and creates competitive pressures as peer organizations respond to similar societal demands. Consequently, enterprises embedded in such environmentally conscious contexts demonstrate a substantially greater willingness to implement sustainable business practices, allocate resources toward green projects, and integrate environmental considerations into core operational decisions. This enhanced commitment to sustainability, in turn, strengthens the positive impact of ESG initiatives on green innovation outcomes by creating alignment between corporate capabilities (reflected in ESG performance) and external incentives (reflected in public attention).

Following the methodological approach established by Wu et al. [12], this research measures the level of public environmental concern using the Baidu Search Index as a proxy variable. This index aggregates search volume data to construct a quantitative indicator of attention intensity. Specifically, "smog" (雾霾) and "environmental pollution" (环境污染) are selected as the primary keywords for identifying and tracking the level of environmental attention, as these terms capture the most salient and widely discussed environmental concerns within the Chinese context during the study period.

The empirical results presented in Table 6 provide compelling support for the hypothesized moderating effect. The interaction term coefficient for public environmental attention registers at 0.00577, achieving statistical significance at the 5% level. This finding carries important implications. It suggests that in regions or periods characterized by higher levels of public environmental attention, the effect of ESG performance on corporate green innovation is improved. In other words, the marginal impact of improving ESG on innovation output is greater when societal environmental consciousness is more pronounced.

Table 6. The moderating effect of Public Attention (PA)

	(1)
	GI
ESG	0.00349***
	(3.095)
ESG x PA	0.00577**
	(2.366)
Size	0.02457***
	(2.587)
Lev	-0.02151
	(-0.479)
Roe	0.12426
	(1.397)
Age	0.03304

	(1.543)
Grow	-0.00651
	(-0.822)
Top1	-0.09634
	(-1.453)
Soe	-0.00767
	(-0.233)
Tan	-0.09402
	(-1.529)
_cons	-0.52208**
	(-2.334)
N	11045
r2	0.74405

6. Conclusion

The empirical evidence presented in this study establishes a robust and statistically significant association between corporate ESG performance and green technology innovation outcomes. Crucially, this relationship exhibits pronounced spatial heterogeneity across China's diverse regional landscape, with enterprises located in the economically developed eastern coastal region deriving the most substantial innovation benefits from their ESG engagement efforts. This geographical disparity likely reflects differences in institutional quality, market maturity, and resource availability that characterize China's uneven development pattern.

Moreover, this analysis reveals that the signalling mechanism through which ESG performance communicates corporate quality and sustainability commitment to external stakeholders among non-state-owned enterprises (non-SOEs). These firms, lacking the implicit government backing enjoyed by state-owned counterparts, gain reputation and resource access advantages from credible ESG signals. Additionally, this research identifies and empirically validates financing constraints as a critical mediating mechanism. Superior ESG facilitates access to capital by reducing information asymmetries, lowering perceived investment risks, and enhancing corporate creditworthiness in the eyes of financial institutions and investors.

Finally, the study demonstrates that environmental public attention works as a significant positive moderator, improving the ESG-innovation relationship when societal environmental consciousness intensifies. These empirical findings reflect several important practical implications for corporate strategy and policy design. Companies across all ownership structures and regions should strategically prioritize ESG practices by systematically integrating environmental, social, and governance principles into their core strategic planning processes rather than treating them as peripheral compliance activities. This recommendation particularly applies to non-state-owned firms confronting severe financing constraints, as enhanced ESG performance offers a viable pathway to alleviating capital access barriers and unlocking innovation potential. Furthermore, enterprises operating in different regional contexts should adopt geographically tailored ESG strategies that account for local institutional characteristics, stakeholder expectations, and development priorities. Eastern enterprises, benefiting from more sophisticated markets and stronger institutional environments, can productively focus on deepening the integration between ESG initiatives and green innovation activities to maximize synergistic benefits. Conversely, firms in central regions should concentrate on foundational capacity building, enhancing internal ESG awareness, establishing systematic ESG management frameworks, and developing organizational capabilities necessary for effective ESG implementation.

Future research can extend this analysis across multiple dimensions. First, research could expand beyond the firm level and explore how ESG collaboration and coordination among supply chain partners generate synergistic impacts on green innovation performance. Second, empirical methodologies could be enriched by leveraging quasi-natural experimental scenarios arising from discrete ESG policy changes or regulatory interventions. This can enable more rigorous causal identification and address potential endogeneity concerns.

References

- [1] Xiao, Q., & Fei, L. (2024). How does climate vulnerability impact green innovation? A hindrance to sustainable development. *Innovation and Green Development*, 3(4), 100169. <https://doi.org/10.1016/j.igd.2024.100169>
- [2] Arefieva, O., Polous, O., Arefiev, S., Tytykalo, V., & Kwilinski, A. (2021). Managing sustainable development by human capital reproduction in the system of company's organizational behavior. In *IOP Conference Series: Earth and Environmental Science* (Vol. 628, No. 1, p. 012039). IOP Publishing.
- [3] Wu, L., Wang, C., Ren, H., & Zhang, W. (2024). How does executive green cognition affect enterprise green technology innovation? The mediating effect of ESG performance. *Heliyon*, 10(14). <https://doi.org/10.1016/j.heliyon.2024.e34287>
- [4] Zhao, C., Wang, Z., Tang, Y., & Yang, F. (2025). ESG performance, green technology innovation, and corporate value: Evidence from industrial listed companies. *Alexandria Engineering Journal*, 123, 369-380. <https://doi.org/10.1016/j.aej.2025.03.097>
- [5] Guo, Z., Wang, Q., Sun, Q., Zhao, H., & Liu, J. (2024). Do companies improve their ability to innovate sustainably as a result of ESG ratings? A quasi-natural experiment based on the ESG rating of SynTao Green Finance. *Total Quality Management & Business Excellence*, 35(11-12), 1183-1211. <https://doi.org/10.1080/14783363.2024.2364196>
- [6] Xiao, J., Zhang, X., Yan, H., & Zeng, P. (2025). ESG performance and green cooperative innovation: a signaling theory framework. *Asia Pacific Journal of Marketing and Logistics*, 1-19. <https://doi.org/10.1108/APJML-04-2025-0612>
- [7] Li, J., Lian, G., & Xu, A. (2023). How do ESG affect the spillover of green innovation among peer firms? Mechanism discussion and performance study. *Journal of Business Research*, 158, 113648. <https://doi.org/10.1016/j.jbusres.2023.113648>
- [8] Zhang, D., & Lucey, B. M. (2022). Sustainable behaviors and firm performance: The role of financial constraints' alleviation. *Economic Analysis and Policy*, 74, 220-233. <https://doi.org/10.1016/j.eap.2022.02.003>
- [9] Kim, J. W., & Park, C. K. (2023). Can ESG performance mitigate information asymmetry? Moderating effect of assurance services. *Applied Economics*, 55(26), 2993-3007. <https://doi.org/10.1080/00036846.2022.2107991>
- [10] Yu, C. H., Wu, X., Zhang, D., Chen, S., & Zhao, J. (2021). Demand for green finance: Resolving financing constraints on green innovation in China. *Energy policy*, 153, 112255. <https://doi.org/10.1016/j.enpol.2021.112255>
- [11] Lin, B., & Ma, R. (2022). How does digital finance influence green technology innovation in China? Evidence from the financing constraints perspective. *Journal of environmental management*, 320, 115833. <https://doi.org/10.1016/j.jenvman.2022.115833>
- [12] Wu L., Yang M. & Sun K. (2022). Impact of public environmental attention on environmental governance of enterprises and local governments. *China Population Resources and Environment*, 32(2):1-14. (in Chinese)

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Research on the Financial Performance of Scenario-Oriented Transformation of Business Models: An Exploratory Case Based on Suning.com

Rui Tang* and Feimei Liao

Jiangxi Normal University, Nanchang, Jiangxi, China

**Corresponding author: Rui Tang*

Abstract

Improving service quality through the scenario-oriented transformation of business models delivers a better service experience for users. Such transformation leads to changes in corporate financial performance, which in turn reflect the effectiveness of the transformation. This paper selects Suning.com Group Co., Ltd. as the research object and analyzes its financial performance from 2017 to 2022 through case analysis and literature research. Firstly, it introduces the research background, significance and methods, then summarizes the relevant theories as the theoretical foundation for the subsequent case analysis. The paper presents an overview of Suning.com, elaborates on the development process of its scenario-oriented transformation of business model, and finally analyzes the impact of this transformation on its financial performance. Through in-depth case analysis, the research conclusions are summarized to provide references for other retail enterprises.

Keywords

business model, scenario-oriented, smart retail, Suning.com

1. Introduction

With the rapid advancement of science and technology and the increasing diversification of consumer demand, traditional business models have gradually failed to adapt to the ever-changing market environment [1]. Therefore, many retail enterprises are seeking an important way for innovative breakthrough and sustainable development—the transformation of business models to be scenario-oriented. This transformation is not only a response to changes in the external environment, but also a necessity for the internal development of enterprises. By building scenarios closely related to consumers' lives, the accurate push of products and services is realized, thereby improving user stickiness and conversion rate [2].

In recent years, the rapid development of technologies such as artificial intelligence and the Internet of Things has provided enterprises with a new perspective and tools, enabling them to gain a deeper understanding of consumer demand and market changes. The application of these technologies has driven the innovation of business models and brought unprecedented development opportunities for enterprises [3].

In the new era, people's increasing requirements for the quality of life have boosted the development of the domestic service economy. The ever-changing economic environment and the gradual maturity of technologies such as big data are driving the progress and development of China's retail industry, which also influences the choice of business models for retail enterprises. To occupy an important position in the retail industry, enterprises must keep pace with the times and enhance their core competitiveness. In the current environment, business model innovation is the most sensible choice [4]. How to innovate business models and improve the economic benefits of enterprises has become the focus of industry development and corporate competition.

2. Case Introduction of Scenario-Oriented Transformation of Suning.com's Business Model

2.1 Overview of Suning.com

Suning.com Group Co., Ltd. (referred to as “Suning.com”), founded on May 15, 1996 and listed on the Shenzhen Stock Exchange in July 2004, is a leading smart retail service provider in China with its headquarters in Nanjing, Jiangsu Province. Its predecessor is Suning Electronics. During the era of Internet retail, Suning.com has continuously promoted the strategies of smart retail and scenario interconnection, expanding its product categories in an all-round way, realizing omni-channel online access and omni-customer integration. By opening up the supply cloud, user cloud, logistics cloud, financial cloud and marketing cloud, it has achieved full coverage from online to offline, from cities to counties and towns, and from shopping malls to communities, providing consumers with an ubiquitous 1-hour scenario life circle solution that comprehensively covers consumers' daily living needs. Its operated commodities include comprehensive categories such as traditional home appliances, department store goods, books, daily necessities and virtual products. In 2020, Suning.com was again listed in the *Fortune* Global 500, and ranked first in the retail industry with a brand value of 296.815 billion yuan in the 2020 *China's 500 Most Valuable Brands*. As of June 2020, Suning.com's offline omni-scenario network has covered the whole country, with various innovative Internet stores of the “one large, two small and multiple specialized” format such as Suning Plaza, Suning Carrefour Community Center, Suning Department Store, Suning Retail Cloud, Suning JIWI and Suning Red Baby, ranking among the forefront of domestic offline networks; Suning.com's online business, through self-operation, open platform and cross-platform operation, has become one of the top players in China's B2C industry, with a total of 602 million retail members.

2.2 The Process of Scenario-Oriented Transformation of Suning.com's Business Model

2.2.1 Online Retail Transformation Stage

After its successful listing in 2004, Suning.com built its own brand and its physical professional retail business was deeply loved by the public. With the development of e-commerce, Suning recognized the future development trend of online retail and started the construction of an online platform. In 2010, Suning.com's online mall was officially launched, entering the e-commerce B2C field and building China's top e-commerce online shopping platform. Due to the rapid development of the Internet, offline retail declined sharply, and Suning.com began to focus on the integrated development of online and offline businesses. To realize retail transformation, Suning.com acquired mother and baby B2C platforms such as “Red Baby” in 2012, which was its first merger and acquisition in the e-commerce field and expanded part of its business categories.

2.2.2 O2O Transformation Stage

The O2O transformation stage witnessed a remarkable development journey of Suning.com. During this stage, Suning.com not only successfully realized the transformation from traditional retail to Internet retail, but also consolidated its position in the fierce market competition, laying a solid foundation for its sustainable development in the future.

In the O2O transformation stage, Suning.com adopted the Internet retail strategy of “one main body with two wings”. The so-called “one main body” refers to Internet retail as the core, deeply integrating the advantageous resources of online and offline businesses. The “two wings” refer to the open platform and smart logistics: attracting more partners through the open platform to jointly build an ecological circle;

improving distribution efficiency and optimizing consumer experience through smart logistics. The implementation of this strategy enabled Suning.com to achieve remarkable results in the O2O field. The company entered the mid-term of new retail transformation in 2013, accelerating the development speed of corporate retail and expanding the strategic layout of corporate retail. In addition, Suning.com changed its corporate name and proposed the Internet retail layout of “one main body with two wings”, achieving success in taking the lead in building the O2O model, implementing the same price for the same products online and offline, and realizing the integration of online and offline businesses. Furthermore, the company acquired PPTV to leverage its online resources for traffic diversion and expand the marketing field. In the O2O transformation stage, Suning.com obtained more high-quality commodity resources by establishing close cooperative relations with suppliers, and also gained greater advantages in terms of price and promotion. It also attached great importance to technological innovation and talent training. In 2014, Suning.com actively promoted the integration of various retail channels and supply chains to ensure the smooth progress of the enterprise's transformation to the O2O model. In the mid-term of new retail transformation, Suning.com adhered to the dual-line integrated development strategy and laid out omni-channel marketing to provide consumers with a one-stop shopping experience.

2.2.3 Smart Retail Transformation Stage

In the smart retail transformation stage, Suning.com deeply recognized the importance of digital transformation. With the rapid development of science and technology and the constant changes in consumer demand, the traditional retail model has been difficult to meet market demand. Through the application of advanced technologies such as big data and artificial intelligence, Suning.com has realized accurate insight into and rapid response to consumer demand. It has also strengthened cooperation with the supply chain, optimized supply chain management, and improved operational efficiency and service quality. Meanwhile, it has enhanced brand building and marketing innovation to attract more consumers' attention and trust, and carried out various marketing activities to stimulate consumers' purchasing desire and boost sales performance. Suning.com also paid special attention to improving user experience by optimizing its shopping process and enhancing after-sales service quality, providing consumers with a more comfortable shopping environment. Its user-centric philosophy made Suning stand out in the fierce competitive environment.

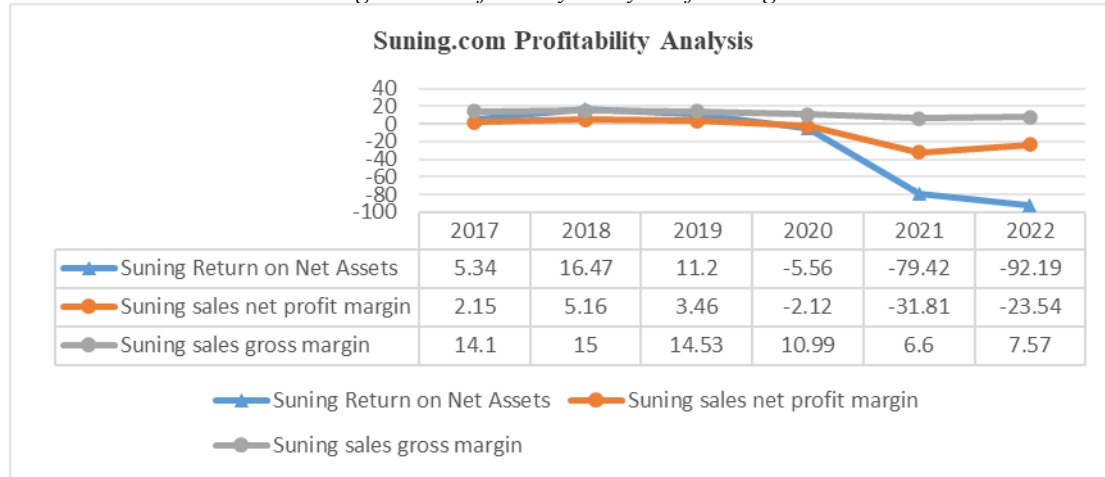
In general, the smart retail transformation stage was a period full of challenges and opportunities for Suning.com. Through digital transformation, online-offline integration, brand building and other measures, Suning.com successfully realized the transformation from traditional retail to smart retail, which enhanced its market competitiveness. Adhering to the business philosophy of innovation-driven and customer-first, Suning.com has continuously promoted the development of smart retail and built a scenario-oriented retail model. Both online platforms and offline stores are committed to providing consumers with an immersive shopping experience and higher-quality services to meet their diversified needs.

3. Analysis of the Financial Performance of Suning.com's Scenario-Oriented Business Model Transformation

3.1 Profitability Analysis

As shown in Figure 1, Suning.com's return on net assets in 2018 was 16.47%, a significant increase compared with 5.34% in 2017, indicating a good profitability performance. It can be seen that the online business has brought financial performance to the enterprise under the influence of scenario-oriented transformation. The return on net assets, net profit margin on sales and gross profit margin on sales have declined year by year from 2019 to 2022, and the return on net assets and net profit margin on sales have been negative since 2020. Since 2019, affected by the COVID-19 pandemic, the income of offline stores has decreased, physical stores have operated in phases and door-to-door installation services have gradually declined, while the proportion of online services has increased. At the same time, the implementation of the expansion strategy to expand market share has led to an increase in operating costs. As a result, the operating income in 2019 was 24.2 billion yuan more than that in 2018, but the net profit in 2019 was 3.32 billion yuan less than that in 2018. The net profit margin on sales also showed a downward trend with the decrease in total sales.

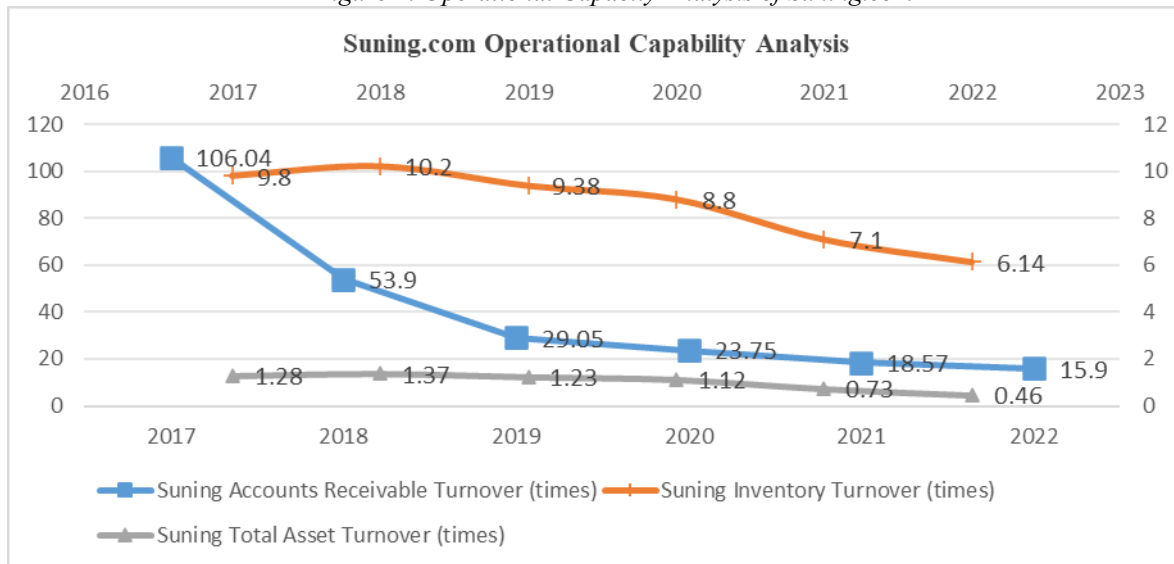
Figure 1: Profitability Analysis of Suning.com



3.2 Operational Capacity Analysis

As shown in Figure 2, Suning.com's accounts receivable turnover rate has been declining continuously, especially a significant drop between 2017 and 2018, falling from 106.04 times in 2017 to 53.90 times in 2018, a decrease of 52.14 times. During the period of business model scenario-oriented transformation, all retail enterprises have shown different downward trends. Due to the transformation of the business model, the expansion of the enterprise scale has directly led to an increase in accounts receivable and payment for goods receivable. In 2019, affected by the COVID-19 pandemic, Suning.com relaxed its accounts receivable policy to ensure transactions with suppliers. This measure reduced Suning.com's accounts receivable turnover rate, but in the long run, it has a promoting effect on the sustainable development of the enterprise and enhances its market competitiveness.

Figure 2: Operational Capacity Analysis of Suning.com

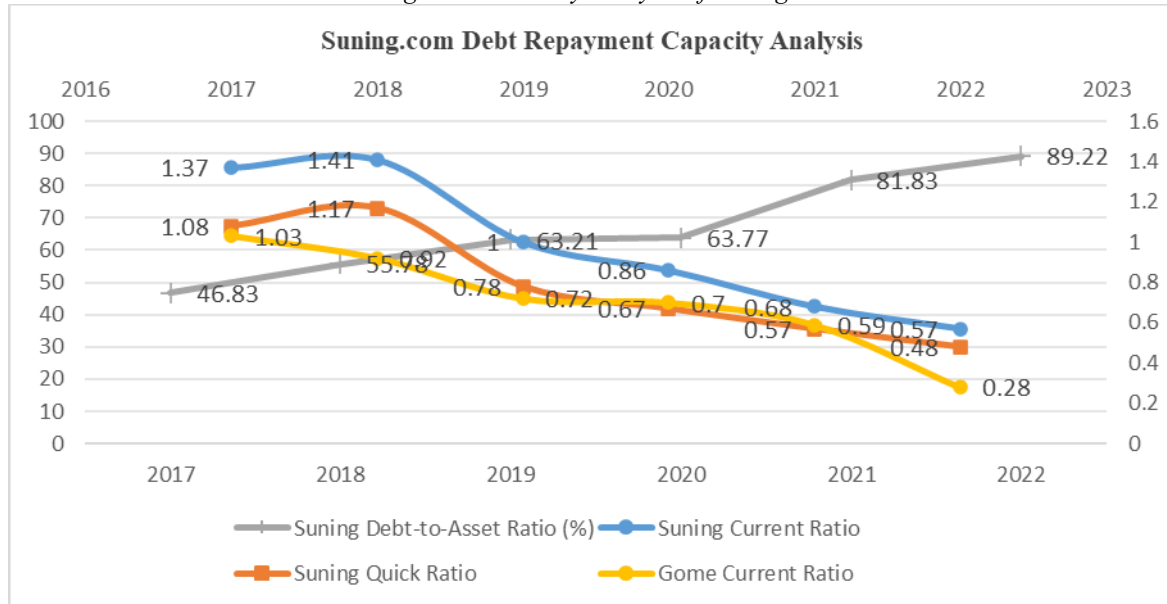


3.3 Solvency Analysis

As shown in Figure 3, Suning.com's asset-liability ratio has been increasing continuously from 2017 to 2022. Although the COVID-19 pandemic broke out at the end of 2019, the asset-liability ratio in 2020 was still 0.56 percentage points higher than that in 2019, with a growth rate of 18.06 percentage points in 2021. Through analysis, it is found that Suning.com issued additional bonds and borrowed loans, leading to an increase in the asset-liability ratio. Current ratio and quick ratio are important indicators to measure an enterprise's short-term solvency. Both Suning.com's current ratio and quick ratio first rose and then fell, increasing from 2017 to 2018, which indicates that Suning.com's short-term solvency and liquidity have been

improved through scenario-oriented transformation. However, they have been declining since 2019, with the most significant drop in 2019. This is not only affected by the COVID-19 pandemic, but also due to Suning.com's acquisition of Carrefour, which generated a large amount of loans and led to a decline in solvency. However, compared with Gome Retail, which did not carry out the transformation, Gome Retail's current ratio has been lower than that of Suning.com and decreased year by year, with the short-term debt repayment pressure increasing year by year. The comparative analysis shows that the scenario-oriented transformation has a certain effect on improving Suning.com's short-term solvency.

Figure 3: Solvency Analysis of Suning.com



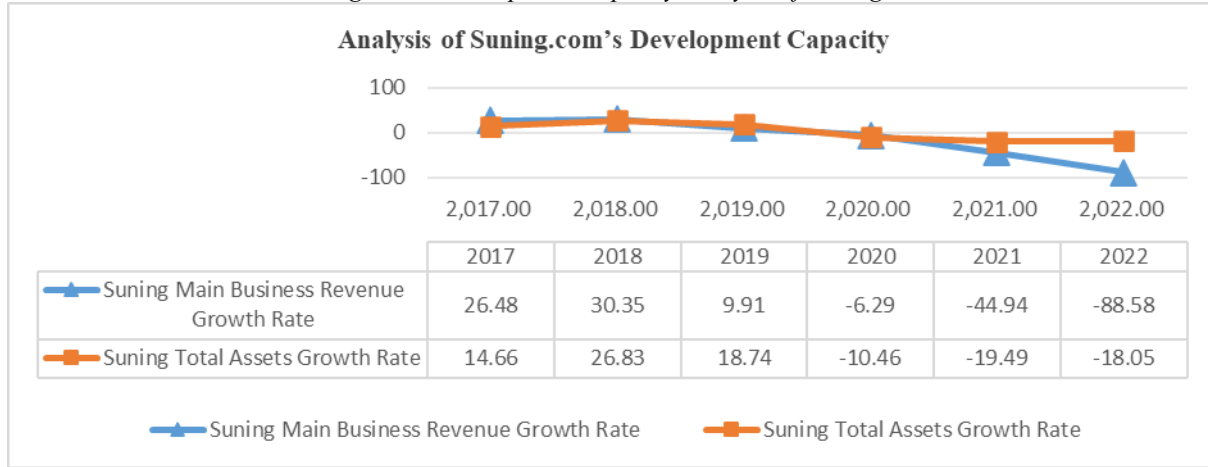
3.4 Development Capacity Analysis

As shown in Figure 4, before the outbreak of the COVID-19 pandemic, Suning.com's total assets were increasing continuously. In 2017, the scenario-oriented transformation of Suning.com led to an upward trend in its accounts receivable, inventory goods and other assets. After the strengthening of O2O operation, the multi-line linkage mode led to an increase in the enterprise's sales volume and a significant growth in main business income. The growth rate of main business income increased from 26.48% to 30.35% between 2017 and 2018, an increase of 3.87 percentage points. Through comparison with other retail enterprises, it is found that the scenario-oriented transformation has improved the enterprise's profitability and brought economic benefits. The growth rate of Suning.com's main business income has been decreasing continuously from 2018 to 2022. The outbreak of the COVID-19 pandemic at the end of 2019 led to a drop in the growth rate of main business income from 30.35% to 9.91%, and it fell to -6.29% in 2020 due to the pandemic, which restricted people's outdoor consumption and the turnover of physical stores. The growth rate of main business income of many retail enterprises has declined, but Suning.com's decline range is relatively small compared with others.

Total asset growth rate, the ratio of net profit to total assets, is an important indicator to measure the speed of a company's scale expansion, reflecting the growth of the company's asset scale in a certain period. Suning.com's total asset growth rate rose from 14.66% in 2017 to 26.83% in 2018, an increase of 12.17 percentage points, indicating the continuous expansion of total assets, the increase in the scale of wealth management products, the expansion of industrial layout and the enrichment of categories. The online-offline integrated model has improved the enterprise's sales scale, demonstrating excellent development capacity. Suning.com's total asset growth rate decreased from 26.83% in 2018 to 18.74% in 2019, because Suning Financial Services and Suning Stores were no longer included in the consolidated financial statements in 2019, resulting in a reduction in asset scale. The total asset growth rate in 2020 was -10.46%. Analysis shows that Suning.com's business decreased due to the impact of the COVID-19 pandemic and the new revenue standard, leading to a reduction in total assets. Relatively speaking, Suning has a good development

capacity, and the significant changes of the enterprise after the scenario-oriented transformation can be seen from the total asset growth rate.

Figure 4: Development Capacity Analysis of Suning.com



4. Conclusions and Recommendations

4.1 Research Conclusions

With the rapid development of the digital economy, traditional retail enterprises have gradually failed to keep up with the pace of development, and many retail enterprises have entered the e-commerce field to meet consumer demand and achieve future development. However, the dividend of e-commerce platforms has become saturated, and many enterprises have to increase costs and face reduced income for survival. Traditional enterprises have begun to adopt the new retail model of integrating online and offline businesses. This paper takes Suning.com as the research object, which integrates various online and offline scenarios and drives consumer shopping through scenario-oriented development. This paper analyzes the impact of the scenario-oriented transformation of business model on financial performance through the analysis of its financial performance, and conducts a horizontal and vertical financial analysis of Suning.com using its data from 2017 to 2022. Through the research, the following conclusions are drawn:

1. Excessively fast expansion speed leads to increased costs

Suning.com is affected by high costs, lacking advantages in the retail industry competition. Blindly expanding the enterprise scale will lead to increased costs without a corresponding growth in operating income.

2. Poor inventory management and low level of accounts receivable management

The total asset turnover rate and inventory turnover rate have been declining continuously, and a large amount of loans has led to a sharp drop in the accounts receivable turnover rate. Suning.com's logistics system needs to be optimized.

3. Unstable development strategy and insufficient risk resistance capacity

The value brought by business model transformation to the enterprise was only reflected in 2018. The subsequent deviation in development direction and strategy has led to poor financial performance of Suning.com and reduced risk resistance capacity.

4.2 Policy Recommendations

Based on the above research conclusions, the following policy recommendations are put forward:

1. Slow down the expansion speed and strengthen cost control

Enterprises should give full play to the advantages of online-offline integration, accelerate the integration of online and offline businesses, optimize the supply chain and realize product differentiation. Reduce

unnecessary offline advertising expenses and make the necessary advertising expenses focus on product promotion rather than just building corporate brand awareness. Close some stores with low profitability to reduce cost expenditure. Suning.com should clarify its positioning and not ignore its capabilities or indicators, which will lead to a decline in corporate competitiveness and is not conducive to its long-term development.

2. Optimize the logistics system and improve inventory management

Suning.com should continue to promote the smart logistics system and improve the enterprise's logistics and warehousing system based on the five elements of price, time, quantity, location and quality. Build a customized logistics system by leveraging digitalization, such as understanding customers' consumption preferences through big data, reaching sales agreements with product suppliers, and producing and processing products according to customers' preferences. With the gradual development of the retail industry, the online market will continue to expand, and Suning.com should expand the distribution scope of logistics, and improve inventory and accounts receivable management.

3. Strengthen internal risk management and adjust development strategies

Suning.com should strengthen internal risk management, continue to promote scenario-oriented construction, realize dual-line integration and build a more comprehensive smart sales model. It should pay attention to the personalized needs of consumers, provide them with higher-quality and more professional services, and improve the economic benefits of the enterprise.

References

- [1] Cai, C. H., Liu, W., Jiang, J. H. The impact of business model scenario orientation on value creation: A longitudinal case study of Rainbow Department Store's digital transformation from 2007 to 2018[J]. Nankai Business Review, 2020, 23(3): 98-108.
- [2] Jiang, J. H. Can “playing to the gallery” create value in business model innovation?—The theoretical origin and creation mechanism of scenario value[J]. R & D Management, 2019, 31(6): 139-154.
- [3] Ming, J. P., Dong, R. Q., Du, Y. Y., et al. Research on the business model of digital transformation in the mining industry[J]. China Collective Economy, 2024(8): 31-34.
- [4] Talaga, J. Online payments in the new retail environment[J]. Journal of the Healthcare Financial Management Association, 2019, 63(3): 86-91.
- [5] Xiong, A. H., Hou, D. E. Research on the path of business model innovation of platform enterprises from the perspective of configuration[J]. Science Research Management, 2022, 43(12): 89-97.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Media Information Dissemination Mechanisms and Capital Market Effects

Haotian Wu*

Beijing-Dublin International College, Beijing University of Technology, Beijing 100124, China

**Corresponding author: Haotian Wu*

Abstract

Driven by digital technology, the linkage between media and the capital market has grown increasingly tight. Traditional media and online new media exhibit markedly different dissemination characteristics: the former is defined by authority, while the latter features viral dissemination accompanied by redundant and false information, along with the “information cocoon” problem. Current academic research lacks a systematic synthesis of the relationship between online new media and the capital market, suffering from issues such as missing key subjects and insufficient timeliness. This study focuses on the three core stakeholders—investors, enterprises, and the government—to explore media information dissemination mechanisms and their capital market effects. It analyzes investors’ information-processing features and behavioral biases in the new media environment, together with the evolution of investment strategies and theories; the recent changes in and influencing factors of corporate social responsibility information disclosure; the challenges faced by the government in regulation, including legislative lag and coordination difficulties, as well as the promotional role of government microblogs. The goal is to provide feasible strategies for these three parties and thereby support the healthy development of the financial market.

Keywords

media information dissemination mechanisms, capital market effects, online new media

1. Introduction

In the current era of rapid digital technology advancement, the connection between media information and the capital market has become ever closer. The efficient operation of the capital market depends heavily on the timely and accurate dissemination of information. As the core carrier of information dissemination, the evolution of media information dissemination mechanisms directly shapes the capital market’s price discovery, resource allocation, and risk transmission.

Mainstream online new media differ sharply from traditional media, which are characterized by specialized information production, single dissemination channels, and strict content review. Their primary dissemination feature is viral dissemination. On social media-driven online new media platforms, anyone can become an information producer, resulting in highly diversified sources yet also abundant redundancy and false information. At the same time, algorithms push content based on user preferences, creating

“information cocoons,” while users actively participate in dissemination through comments, forwards, likes, and other interactions, accelerating the viral spread of information [1, 2].

These differences in dissemination mechanisms have profoundly altered both the mode and intensity of media information’s impact on the capital market. The authority of traditional media makes their content more readily accepted as an “official signal”; for the capital market, especially financial media, their influence on the economic characteristics of the era and on market dynamics tends to be long-term and stable. In contrast, the immediacy and interactivity of online new media position them as “amplifiers” of short-term market sentiment, which may intensify market volatility and even trigger financial risks [2]. Therefore, a deep analysis of the relationship between media information dissemination mechanisms and capital market effects is of great significance for accurately assessing financial risks and maintaining market stability.

Building on the above research background, the investigation of media information dissemination mechanisms and capital market effects holds important practical and theoretical value for the three key stakeholders: investors, corporate entities, and the national government. Existing literature on the relationship between online new media and the capital market lacks systematic organization. On the one hand, research foci are often narrow; for example, studies by Wu Jinyan et al. start from financial media and omit online new media as a subject. On the other hand, temporal gaps are large, and the timeliness of research such as that by Zheng Tao et al. has weakened. Given the rapid development of new media in recent years and the diversification of the media industry, periodic reflections are necessary [3, 4]. Accordingly, the present study centers on the three stakeholders—investors, corporate entities, and the national government.

For investors, the new media environment features information explosion and the difficulty of distinguishing truth from falsehood; “information cocoons” and “emotional contagion” easily lead to blind decision-making. By studying media information dissemination mechanisms, investors can clearly understand the characteristics of different media outlets, learn to assess the reliability of information sources, avoid being misled by false information or extreme emotions, and thus make more rational investment decisions and reduce losses caused by herd behavior [5-7]. For corporate entities, the broad reach of new media imposes higher demands on the transparency and authenticity of information disclosure. Research on media dissemination mechanisms can help enterprises recognize that any false disclosure or concealment may rapidly ferment online, triggering market trust crises and sharp stock-price declines [6, 8]. For the national government, the complexity of new media increases the difficulty of financial regulation, making issues such as false information dissemination and market manipulation harder to control. By clarifying the interactive patterns between media information dissemination and the capital market, the government can refine its regulatory mechanisms in a targeted manner [7].

Drawing on prior literature, this study provides a phased summary of the relationship between media information dissemination mechanisms and capital market effects. It focuses on analyzing the impact of new media on different capital market stakeholders and proposes feasible strategies for investors, corporate entities, and the national government.

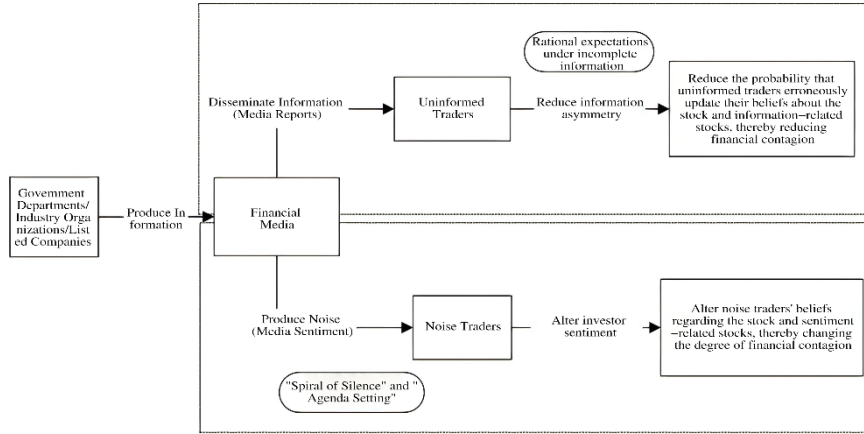
2. Definitions and Differences Between Traditional Media and Online New Media

Traditional media refer to the primary forms of mass communication used before the internet era, including newspapers, television, radio, and magazines. Prior to the internet age, traditional media played a crucial role in disseminating news, information, and entertainment content. Their main characteristics are one-way transmission, slow dissemination speed, and high quality.

Compared with traditional media, online new media exhibit features such as information fragmentation, complexity, timeliness, broad capacity, freedom, openness, bias, and interactivity [1, 9]. These characteristics have introduced various social issues in information dissemination, including threats to social stability and challenges to government regulation. The present study focuses specifically on the capital market segment within society.

Among traditional media with significant relevance to the capital market, financial media take the lead. Wu Jinyan et al., based on financial media, investigated the stock-dominated capital market and found that financial media influence investor sentiment in the capital market by transmitting information and amplifying emotions, thereby reducing the likelihood of financial contagion (see Figure 1) [3].

Figure 1: Influence Pathways of Financial Media on Traders



In the new media environment, capital market participants face information saturation. Gan Qinyu et al. argue that AI large language models represented by ChatGPT can efficiently generate and process vast amounts of information, making it particularly difficult to identify false “noise” within that information. The shift from the limited information of traditional media to the excessive volume of online new media leads to investor information overload, which hinders the smooth incorporation of information into prices and impedes the normal functioning of various market mechanisms [10]. Wang Rongrong et al. contend that internet media effects are complex and lack singularity or predictability in their impact on the capital market; they propose agenda-setting theory and reputation theory [5].

According to data from the 56th Statistical Report on China’s Internet Development released by CNNIC in June 2025 (see Figure 2), the scale of Chinese internet users has reached 1.123 billion, and mobile internet users have reached 1.116 billion. This highlights a pronounced trend toward universal internet adoption. In terms of website and IP domain name counts, internet influence continues to grow daily. This is reflected in agenda-setting theory and reputation theory through greater media influence on “issues” and stronger effects on corporate reputation.

Table 1: Key Indicators from the 56th Statistical Report on Internet Development (June 2025)

Internet users	1.123 billion
Mobile internet users	1.116 billion
IPv6 addresses (/32)	68,567
IPv4 addresses	343.19 million
Total domain names	32.62 million
.CN domain names	20.85 million
Total websites	4.56 million

3. Investor Behavior

From the perspective of investor behavior theory, Wang Rongrong et al. propose two behavioral economics assumptions: (1) investors generally do not process information correctly; and (2) even when investors correctly interpret given information, their final decision-making behavior is not necessarily optimal. Through analysis of investor behavior, they identify four dimensions: information processing, behavioral biases, investment strategies, and investment theories [5].

3.1 Information Processing

- 1) Intuitive judgment: When processing the massive volume of internet information, investors unconsciously treat information as fact due to over-reliance on intuition.
- 2) Overconfidence: The “intellectual arrogance” phenomenon manifests in the new media era as investors believing they already knew internet reports in advance, leading to irrational investment behavior.

- 3) Confirmation bias: Investors seek information consistent with their own views; in the internet era, this appears as the “information cocoon.”
- 4) Conservatism: As age increases and population aging intensifies, conservatism gradually strengthens, manifesting in the capital market as slower reactions to asset price changes.

3.2 Behavioral Biases

- 1) Framing bias: When facing identical information, investors experience cognitive bias due to different frames used by the media, which affects their behavior.
- 2) Herding effect: Investors assume others’ information is correct and allow others’ information and behavior to influence their own judgment, generating irrational investment sentiment and following trends in the capital market.
- 3) Disposition effect: As an anomaly in behavioral finance, this refers to investors’ tendency to sell stocks when prices rise and continue holding them when prices fall.
- 4) Mental accounting: Investors hold different psychological expectations and understandings for the same resource, resulting in different investment behaviors toward the same asset.

3.3 Investment Strategies

Different underlying assumptions lead to two trading strategies:

- 1) Momentum trading strategy: When media release positive news that drives stock prices upward, investors capitalize on upward momentum by buying before the price reaches its peak, thereby earning returns.
- 2) Contrarian trading strategy: Stocks that have risen for several consecutive days tend to fall back to their fundamental value, while stocks that have declined continuously tend to rebound later. Contrarian trading involves buying underperforming stocks.

3.4 Investment Theories

The development of investment theory can be divided into three main stages. Early investment theory (before 1950) laid the foundations for modern investment theories, including the efficient market hypothesis, prospect theory, and dynamic investment strategy theory. In the classical investment theory stage (1950-1980), Markowitz proposed portfolio selection theory in 1952; Sharpe, Lintner, and Mossin developed the Capital Asset Pricing Model (CAPM) between 1964 and 1966; and in 1970, Eugene Fama proposed the efficient market hypothesis, arguing that asset prices already reflect all available information. Markets are classified into three forms—weak-form, semi-strong-form, and strong-form efficient—to distinguish degrees of information reflection. Since the 1980s, improvements to classical investment theory have been modest, with research focus shifting toward breaking the “investor rationality” assumption. Kahneman and Tversky introduced prospect theory, pointing out that investors exhibit irrational behavior and that losses cause greater pain than equivalent gains bring pleasure, thus challenging the premise that “all investors are rational.” Building on prospect theory, behavioral finance theory emerged, incorporating psychological and sociological factors and proposing the Behavioral Asset Pricing Model (BAPM). This model distinguishes “informed traders” from “noise traders” and emphasizes the significant market impact of large-scale “noise.” Its core difference from traditional theory is that it no longer assumes complete investor rationality but instead examines how irrational behavior affects market prices and returns.

In the context of social media, Ding Hui et al. (2018) found that improved investor information capability under social media conditions can significantly reduce stock price crash risk, indicating that enhancing investor information capability is an effective way to lower the risk of listed companies’ stock price crashes [7]. Cai Guilong et al. (2022) discovered that stronger investor–company information interaction helps reduce firms’ cost of equity capital, with a stronger effect for companies with higher information asymmetry, lower institutional ownership, and larger retail investor trading volumes [11]. Lu Rui et al. (2023) found that investor interactions in stock forums amplify the impact of corporate rumors on stock prices, particularly for firms with higher information asymmetry and lower institutional ownership. Overall, investor interactions on

social media may accelerate the spread of corporate rumors and amplify investor sentiment, which is detrimental to the stable development of the capital market.

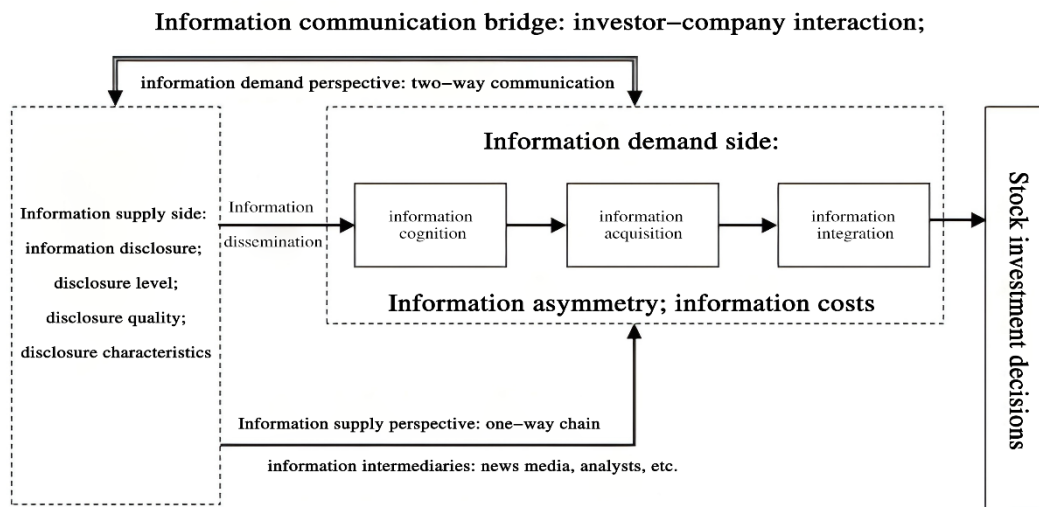
In summary, the likelihood of the four existing dimensions occurring increases significantly in the context of massive information volumes. Investors therefore need to reduce information asymmetry while processing and handling information, minimize biases arising during information processing, adopt appropriate investment strategies, and lower the probability of irrational behavior.

4. Corporate Behavior

The corporate social responsibility (CSR) information disclosure system originated in Western countries in the last century. It is defined as the overall quality level of externally published reports that include information on the assumption and fulfillment of social responsibilities, specifically reflected in the authenticity and reliability, completeness and detail, and vertical/horizontal comparability of the disclosed CSR-related information. In the early stages of the development and improvement of corporate responsibility information disclosure, Western researchers such as Howard proposed that CSR is the primary responsibility of corporate managers, while Mathews argued that enterprises should attach equal importance to social responsibility information and financial information. By the 21st century, scholars such as Tynes and Laroche found that CSR disclosure brings outstanding financial performance and that a bidirectional relationship exists between CSR information and financial performance. Building on this, Dong Yiren et al. discovered that through CSR reports, enterprises can demonstrate their operational and management strength; such reports also help build brand effects and enhance corporate value [12], once again confirming the importance of CSR information disclosure.

In the new media environment, Blankespoor et al. (2014) were the first to reveal the role of social media platforms such as Twitter in information dissemination. They found that technology companies with low market attention can reduce information asymmetry by using Twitter to release and disseminate financial report information, enabling the market to react more fully and reasonably to their financial information [13]. Subsequently, scholars have paid increasing attention to the impact of social media on corporate behavior. Dou Chao and Luo Jinbo (2020) proposed that platform interactions between minority shareholders and companies significantly enhance the sensitivity of executive compensation to performance, with the relationship being more pronounced in firms with higher investor protection, stronger industry competition, and fewer Type II agency problems [14]. Building on this, Cai Guilong et al. (2022), using the investor-listed company interaction platform (a form of information disclosure, see Figure 2), verified three hypotheses: (1) the investor-listed company interaction platform reduces the cost for investors to acquire and integrate corporate information; (2) it improves the quality of information acquisition and integration; and (3) it also reduces information asymmetry among investors [11].

Figure 2: Information Dissemination Chain in the Capital Market



Chang Yingying (2025) examined whether executives' media background helps improve corporate information disclosure quality. Using methods such as variable and model substitution, fixed individual effects, propensity score matching (PSM), instrumental variable approach, and the Heckman two-stage method to control for endogeneity, she found that executives' media background has an inhibitory effect on corporate ESG information disclosure [12].

In summary, the forms of corporate information disclosure in the new media context are no longer limited. By developing efficient disclosure models, enterprises can better adapt to the massive information flows of the new media environment and improve the quality of ESG disclosure. The information dissemination chain in the capital market (Figure 2) also shows that, under the new media background, corporate information disclosure involves diverse intermediaries (news media, analysts, etc.) that can reduce information costs and alleviate information asymmetry, thereby enabling enterprises to better support clients' investment decisions.

5. Government Behavior

In 2025, amid the vigorous development of new media, artificial intelligence, and big data, information dissemination patterns have undergone dramatic changes, presenting numerous new challenges for government efforts to build a sound capital market. A thorough examination of these issues is of great significance for promoting the stable and efficient development of the capital market. Drawing on Wu Guangzhi's 2012 research and Yu Tingting's 2022 case analysis of the "Bank of China Crude Oil Treasure" incident [7, 8], the following key problems are identified:

5.1 Legislative Pace Lags Behind New Media Iteration

New media platforms (such as Weibo and Douyin) develop rapidly with distinctive dissemination characteristics, yet capital market legislation has not kept pace. Current laws lack clear provisions for emerging services such as short-video stock recommendations and live-streaming investment strategies, allowing lawbreakers to engage in illegal operations via new media, disrupt the market, and harm investor interests.

5.2 Lack of Unified Coordination in Regulatory Regulations

Multiple departments issue their own regulatory rules, resulting in a complex system with differing or even conflicting standards and procedures. For example, market regulators apply the Advertising Law to financial product advertisements, while securities regulators impose requirements from a financial perspective, creating operational obstacles for enterprises and reducing regulatory coordination efficiency.

5.3 Coordination Difficulties Among Regulatory Bodies

First, overlapping responsibilities lead to both regulatory gaps and duplication. Agencies such as cyberspace authorities, financial regulators, and industry self-regulatory organizations each emphasize different aspects of new media financial information platform oversight; complex issues are easily shuffled, creating blind spots and difficulties in assigning accountability. Second, information sharing is inadequate: independent information systems, inconsistent data standards, and imperfect sharing mechanisms make it difficult to integrate data for real-time monitoring, resulting in slow responses to violations. When public opinion events occur, departmental barriers are evident and joint action motivation is weak.

5.4 Shortcomings in New Media Awareness and Crisis Response

Some regulators have limited understanding of new media's viral, decentralized dissemination characteristics. Relying on traditional management mindsets easily leads to information omissions and delays in formulating effective strategies. False information on social media can readily trigger stock price volatility; without timely response, market panic may intensify. Failure to fully utilize new media for collecting public feedback and reports, or to follow up promptly on company issues exposed on platforms, results in inadequate responses to public demands and severely undermines government credibility and public supervision enthusiasm.

5.5 Deficiencies in New Media Control and Response Capabilities

Technological tools lag behind: public opinion monitoring systems can only capture simple keywords and struggle with semantic or sentiment analysis, missing critical information and failing to detect sensitive content on niche platforms. There is also a lack of professional analytical models and teams, making it difficult to assess public opinion trends and resulting in passive responses.

Interaction mechanisms are incomplete: policy formulation does not adequately solicit opinions from platforms and the public, inviting criticism; crisis handling lacks close collaboration with platforms, hindering timely release of authoritative information to counter rumors; and communication with the public remains largely one-way, with ineffective feedback collection, which impairs management and guidance.

Promotional role of government microblogs: Xuan Yuhao's 2022 research shows that although local government microblogs in China exhibit relatively scattered temporal and spatial distribution during operation, they perform prominently in information dissemination scale and posting frequency. The content they release not only maintains high textual quality but also centers on positive and constructive emotional tones, primarily focusing on local livelihood protection, social governance, and political developments. The study also finds that government-level information publicity work can further optimize local market operating mechanisms by enhancing regional social trust—specifically, by reducing investors' doubts and questions about local enterprises and significantly improving the market valuation and earnings response coefficients of local companies' earnings announcements [14].

As an important platform for government publicity, government microblogs leverage Sina Weibo to enhance dissemination effectiveness. They offer wide coverage and distinctive media traits such as open dissemination and strong public opinion guidance, all of which strongly support improved government publicity efficiency. From a core value perspective, government microblogs function in four main ways: (1) bidirectional empowerment—enabling low-cost, real-time delivery of policy content (e.g., epidemic prevention deployments) while creating more channels for citizens to participate in political affairs; (2) trust building—through people-oriented publicity that enhances the public's sense of closeness to the government; (3) resource integration—local governments can coordinate information resources across departments and concentrate releases on livelihood hotspots, effectively lowering the public's information acquisition costs (by 2020, the total number of government microblogs had exceeded 100,000); and (4) innovative practice—tailoring differentiated publicity formats to regional characteristics to strengthen guidance over social opinion. Overall, government microblogs, with their authority, timeliness, and interactivity, have reshaped the communication model between government and citizens and become a vital carrier in the public service domain.

Additionally, statistics from Li Yi et al. (see Table 1) indicate that among listed companies in various industries, the education sector has the highest proportion of official Weibo accounts at 75%—though this may be influenced by the small total number of listed companies in the sector. Overall, listed companies in the tertiary sector with high technological content and the need for frequent user interaction show stronger willingness to open Weibo accounts; Weibo applications have now penetrated every industry. To further enhance government competitiveness, it remains necessary to continuously strengthen governmental influence in the new media era.

Table 2: Weibo Account Distribution by Industry

(1) Industry	(2) Accounts Opened	(4) Total Companies	(5) Proportion (%)
Agriculture, Forestry, Animal Husbandry & Fishery	11	44	25.000
Mining	8	78	10.256
Manufacturing	649	2647	24.518
Electricity, Heat, Gas & Water Production & Supply	11	114	9.649
Construction	25	101	24.752
Wholesale & Retail	59	168	35.119
Transportation, Warehousing & Postal Services	23	106	21.698
Accommodation & Catering	5	10	50.000
Information Transmission, Software & IT Services	150	336	44.643
Finance	51	122	41.803

(1) Industry	(2) Accounts Opened	(4) Total Companies	(5) Proportion (%)
Real Estate	32	124	25.806
Leasing & Business Services	24	57	42.105
Scientific Research & Technical Services	12	58	20.690
Water Conservancy, Environment & Public Facility Management	26	72	36.111
Education	6	8	75.000
Health & Social Work	6	12	50.000
Culture, Sports & Entertainment	29	59	49.153
Conglomerates	0	16	0.000
Total	1127	4132	27.275

In summary, as an authoritative body, the government must fully leverage its authoritative position in the new media environment by promptly issuing government microblogs to stabilize market sentiment and reduce the likelihood of irrational behavior. It should also improve relevant laws and regulations, enable efficient coordination among financial management departments with clearly defined responsibilities, and better fulfill the government's role as a "leader."

6. Conclusion

6.1 Research Conclusions

This study examines media information dissemination mechanisms and their capital market effects, focusing on the three key stakeholders—investors, enterprises, and government—and reaches the following conclusions. At the level of stakeholder characteristics and behavior, the fragmentation and interactivity of online new media have intensified investors' information overload and "information cocoon" problems, making them prone to biases such as intuitive judgment and overconfidence during information processing; social media interactions may also amplify the impact of rumors on stock prices. Corporate information disclosure forms have become more diverse in the new media context; investor-listed company interaction platforms can reduce information costs, yet executives' media backgrounds inhibit the quality of corporate ESG disclosure. Governments face challenges including legislative lag behind new media iteration, insufficient regulatory coordination, and inadequate crisis response capabilities, while government microblogs can optimize market mechanisms by enhancing social trust and thereby help stabilize the capital market. At the practical value level, the study clarifies the differentiated impacts of different media types on the capital market and provides targeted action directions for the three stakeholders, which is of great significance for alleviating information asymmetry and promoting the healthy development of the financial market.

6.2 Research Limitations

First, data limitations exist: some literature data lack timeliness, and coverage of dynamic capital market data in the new media environment is insufficient, which may affect the real-time applicability of the conclusions. Second, the research scope is limited to the domestic capital market and lacks comparative analysis of media dissemination mechanism differences across different institutional environments, preventing a full revelation of how institutional factors moderate the "media-market" interaction. Third, platform segmentation is inadequate; the study does not deeply analyze dissemination characteristic differences among platforms such as Weibo, Douyin, and AI-generated content platforms, nor their heterogeneous effects on capital market outcomes. Fourth, the influence of third parties on the capital market—such as financial analysts and other specific individuals—is not considered. Future research can expand in these four directions.

6.3 Future Prospects

Future research can advance in three aspects: First, expand data dimensions by combining real-time capital market trading data with new media platform information dissemination data to conduct cross-period, large-sample studies, thereby enhancing the timeliness and accuracy of conclusions. Second, refine research

perspectives by deeply analyzing differences in media influence across industries, firm sizes, and regional investors, and explore industry- and region-specific patterns of media information dissemination. Third, enrich research methods by introducing empirical analyses—such as constructing econometric models to quantify the impact of media information dissemination on stock price volatility and corporate performance—or conducting case studies of typical new media capital market events, thereby providing stronger empirical support for the interactive relationship between media and the capital market and further improving the relevant theoretical system and practical guidance.

In conclusion, the interaction between new media and the capital market is a complex system shaped jointly by technology, institutions, and human nature. This study's phased summary not only offers practical reference for current market participants and regulators but also reveals the core logic: only by following the principle of “adaptation” and enabling media dissemination mechanisms to serve the dual goals of “price discovery” and “risk prevention and control” in the capital market can the financial market achieve more efficient and robust development in the digital era and provide solid support for resource allocation in the real economy.

References

- [1] Sun, Z., Reflections on integrated development strategies of traditional media and new media. *Caixie Bian*, 2025 (6):75-77.
- [2] Chen, H. Y., Research on problems and countermeasures of online rumor information governance in the new media environment. *News Research Guide*, 2025, 16 (14):115-119.
- [3] Wu J. Y., Research on the impact of financial media on financial contagion: taking the stock market as an example. Dissertation. Southwestern University of Finance and Economics, 2021.
- [4] Zheng, T., Media reporting and capital market development: empirical evidence from Chinese listed companies. Dissertation. Southwestern University of Finance and Economics, 2010.
- [5] Wang, R. R., Research on the impact of internet media on investor behavior in the capital market. Dissertation. Anhui University, 2018.
- [6] Li, Y., The impact of social media on the capital market and its participants. Dissertation, Tianjin University, 2022.
- [7] Ding, H., Lü, C. J. and Huang, H. J., Social media, investor information acquisition and interpretation ability, and earnings expectations: evidence from the “SSE e-Interaction” platform. *Economic Research Journal*, 2018, 53 (1):153-168.
- [8] Yu, T. T., Research on coping with sudden financial network public opinion crises from a new media perspective. Dissertation. Chongqing Technology and Business University, 2022. <https://doi.org/10.27713/d.cnki.gcqgs.2022.000726>
- [9] Wu, G. Z., Information dissemination and government regulation in the new media context. Dissertation. Minzu University of China, 2012. <https://doi.org/10.7666/d.y2137235>
- [10] Gan, Q. Y. and Luo, R. H., Media reporting and capital market pricing efficiency from the perspective of information overload. *Securities Market Herald*, 2025 (8):66-79.
- [11] Cai, G. L., Zhang, Y. N. and Xu, Y., Investor-listed company interaction and capital market resource allocation efficiency: empirical evidence based on the cost of equity capital. *Management World*, 2022, 38 (8):199-217. <https://doi.org/10.19744/j.cnki.11-1235/f.2022.0118>
- [12] Dong, Y. R. and Zhao, S. M., The impact of corporate social responsibility in different domains on consumer responses: an empirical study based on private car buyers. *Xuehai*, 2010 (5):148-153. <https://doi.org/10.16091/j.cnki.cn32-1308/c.2010.05.036>
- [13] BLANKESPOOR E, MILLER G S, WHITE H D. The role of dissemination in market liquidity: Evidence from firms' use of Twitter™. *The Accounting Review*, 2014, 89(1):79-112.

- [14] Dou, C. and Luo, J. B., Can minority shareholders improve executive compensation contracts by “voicing” through social media. *Finance & Trade Economics*, 2020, 41 (12):85-100. <https://doi.org/10.19795/j.cnki.cn11-1166/f.20201214.008>
- [15] Chang, Y. Y. and Chen, X. W. and Lin, X. R., Does executives’ media background help improve corporate ESG information disclosure quality?. *Journal of Nanjing Audit University*, 2025, 22 (2):45-54.
- [16] Xuan, Y. H., Government information publicity, social trust, and decision-making by corporate managers and investors. Dissertation. Southwestern University of Finance and Economics, 2022. <https://doi.org/10.27412/d.cnki.gxncu.2022.000359>

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Application of Generative Artificial Intelligence in Financial Innovation and Risk Management

Jianheng Zhang*

School of Finance, Zhejiang University of Finance & Economics, Hangzhou, China

**Corresponding author: Jianheng Zhang*

Abstract

In the age of burgeoning technology, AI is playing an extremely indispensable role in finance. This paper focuses on the recent hot topic of two novel applications of generative AI, robo-advisors and emotionally-aware AI service, and the following risk prevention and control of them. And the thesis reviews the previous studies from 2022 to 2024 selected on AI core technology, machine learning, AIGC, robo-advisors and risk prevention. The review finds that during the growth of AI, though generative AI significantly improves the working efficiency, the methods to manipulate generative AI legally and avoid extra ethical and safety problems remain to be solved. Hence, the review proposes future research directions on AI explainable framework and regulatory prevention through human intervention to figure out the problems scientifically. Furthermore, this paper provides a holistic reference for both researchers and employees hoping to enhance and perfect generative AI, and address the security issues that are still not resolved at present.

Keywords

AIGC, robo-advisor, risk prevention

1. Introduction

Based on the big data, globally, almost 50% investors refer to generative AI when they are buying financial products, and more than 70% companies dispose financial business segment with the help of generative AI. Generative AI is a branch of AI that can output text, picture, audio and so on in a short time according to the user's need [1], which brings about production efficiency, new track creation and technological inclusion. The success and widespread of generative AI are credited to technological breakthroughs, algorithm support, updated data, an open source ecosystem and research and development sources. Generative AI has evolved into an extraordinarily sophisticated and mature field in recent years.

Artificial Intelligence Generated Content (AIGC) technology becomes the core technology of content production, and it is broadly applied in GPT-4, Stable Diffusion, Sora, Github Copilot and so on. In the field of finance, it can improve the university and enable people to enjoy personal service in investment and consultation. AIGC push generative AI from theory to practice, employing a human-machine collaboration operation mode and structured data output to enable generative AI to better address human needs. In finance, it can also automatically generate and review contracts, process efficiently and output display of data.

As generative AI continues to advance and innovate, it plays an increasingly significant role in finance. Generative AI has gradually become the business core in the financial industry, leading in services, operations, etc. The financial industry prioritizes efficiency and risk detection. In trading, there is an extreme requirement for speed. Even a difference of a few milliseconds can make a huge difference. Generative AI needs to leverage its decision-making by precisely reacting to users' needs in the shortest possible time. However, risk prevention is becoming increasingly significant. Detecting and preventing financial fraud [2], formulating current AI regulations and governance [3], mapping the ethical and legal risks of generative AI in the financial sector [4], and identifying the potential challenges and limitations [5] are the focuses of current research. Therefore, delving into the working principle of generative AI and preparing for potential risks proves necessary. Due to the excessively rapid advancement of AI, pressing methodology and practice time are urgently needed to guide AI.

This paper teases financial achievements with the contribution of generative AI in many aspects and supplements. Moreover, the paper mainly reviews the rising robo-advisors about their current situation and future expectations, and the following efficient operations and production innovations brought by generative AI, with the rising risk issues due to the rapid development of technology. Therefore, this paper also proposes risk prevention measures and regulatory measures for generative AI, aiming to summarize frontier defense methods for both regulators and investors in the era of developing the internet.

2. Overview of Generative AI in Finance

2.1 The Technical Core of Generative AI

The field of generative AI can be traced back to 2018 when the first AI model, ChatGPT-1, was created by OpenAI. It aroused heated discussion and concerns at that time. AI refers to a specific capability for machines of thinking, learning and make decisions like humans. Based on massive sample learning and data processing, the core operational logic of the GPT (Generative Pre-trained Transformer) series, GPT, has become a benchmark of generative AI. By 2026, the GPT series will have released the latest version GPT-5, symbolizing that the generative AI system has matured in technique. Compared to conventional AI, generative AI is a completely novel type of AI that can create entirely new content based on the Variational Auto-encoder (VAE) and Generative Adversarial Network (GAN) models [6]. Through the VAE algorithm, generative AI extracts information, decodes it, and then outputs something similar to the original data. Combined with GAN, another generative model that trains its generator and discriminator to compete and learn from each other, generative AI collects and identifies massive data based on the Large Language Model (LLM), thereby producing valuable and precise information. These emerging technologies enable generative AI to make breakthrough on comprehending knowledge and creating corresponding solutions to problems that never occurred before.

How generative AI masters these technologies is listed as follows. From two types of study, machine learning (ML) and deep learning (DL), generative AI is empowered by intelligence. Machine learning depends on human guidance, take images as an example, when feeding enough trading trend charts to AI, it can identify the trend of rise and fall, and recognize similar historical trends. Deep learning requires more computational load and data volume, and it relies more on AI itself. In deep learning, discriminative study, honored as a classification master, can distinguish and judge information, while a generative study can create diversified, reasonable suggestions in various fields. AIGC takes great effect in one of the most important segments of content creation, and takes the top priority in the output of generative AI. After sorting out the characteristics of certain items, AIGC generates totally new texts, images, videos, etc. by following the requirements and memory after learning. For computers, their memory is long-lasting, which helps them reserve and integrate multi-dimensional data across multiple timelines. And it can leverage AI maximized stored knowledge, this is friendly for humans to acquire knowledge.

AIGC has a huge power of content generation, code compilation, and complex reasoning. In terms of generating, there are many instruction types like text-to-text generation, text-to-image generation, text-to-video generation, and each one matches its own model. For instance, a typical type of text-to-text generative AI ChatGPT and DeepSeek rely on LLM, DouBao and Sora depend on a video generation model. The principle of large models is mainly divided into three parts: training, inputting, and generating, respectively

corresponding to summarizing language patterns, dividing the questions into tokens and predicting the answers word by word.

2.2 The Theoretical Logic of AIGC Empowering the Financial Industry

There is an even closer trend in the integration of generative AI and finance. So, many new financial applications have emerged in recent years, which brings vitality to the financial field. Generative AI is getting familiar with basic financial phenomena and the corresponding coping approaches through continuous machine learning, including advanced theories, background recognition, model implementation, transformers, core network framework structures, and so on. With the popularization of the Internet, more and more people prefer to refer to AI suggestions when they are trading or buying financial products. Undoubtedly, they can obtain comprehensive and complete information from AI and, to a large extent, get rid of emotional factors.

And it keeps AI closely related to current affairs and hot topics, which is beneficial for AI to obtain the latest information in case of information lag. It is undoubtedly friendly for institutions and individuals because of information equality and symmetry.

AIGC has an absolute advantage in information acquisition. AI can instantaneously scan and process a large amount of heterogeneous information, while identifying easily overlooked patterns and connections. Without human subjective thoughts, AIGC enhances objectivity and breaks down the information silos, boosting cross-disciplinary integration. Key functions include structuring unstructured data, integrating multilingual information seamlessly and understanding the implicit requirements. Vocabulary strategy prompts containing explicit frame tasks, export constraints and citation traceability awareness are effective and essential for leveraging AIGC. For the general public, it is universally beneficial because AI breaks the limitations of language and region.

Finance is the domain where AIGC excels. The financial industry is closely linked with mathematics, knowledge and decision-making, and a vast amount of structured data and unstructured data provide AIGC with an extremely awesome study environment. AIGC can break down statistic silos and integrate and understand data across different domains, which helps financial professionals generate financial, operational and customer profiling data. Afterwards, it can automatically organize into reports and provide investment and asset allocation recommendations based on the report. Then, it will check for corrections in a short time and identify the potential risk in the report. In addition, the inclusiveness of AIGC has brought increasing convenience to people and made it more practical.

3. The Cutting-edge Appliances of Generative AI in Finance

3.1 Intelligent Interaction and Marketing of Generative AI

Credited to large language models (LLMs) and databases, generative AI masters many practical financial applications. As proven by Buffett's 10-year experiment, passive investment always performs better than active investment. Passive investment vehicles track markets by holding portfolios that mirror representative indices, and the emerging investment pattern, the robo-advisor, is perfectly following this trend.

Contrary to the stereotype, the robo-advisor breaks through the limitations of the traditional model, injecting vitality into self-driven wealth management, which brings commendable personalized service and adaptability. The robo-advisor also supports multiple investment methods, including stocks, bonds, funds, exchange-traded funds (ETFs), real-world assets (RWAs), private equity, commodities, and indices, meeting different users' needs. For example, in terms of those under 20, between 25-40, and above 60, the robo-advisors recommends different configurations for diverse risk appetites and investment proportions. Generally, the older the investor is, the less risk they should assume. The robo-advisor captures users' risk appetites, follows investment theories, and, according to mathematical algorithms, precisely formulates a scheme for users [7]. The robo-advisor helps investors in various ways, such as determining the investment proportion, automatically and ideally allocating investors' assets, and calculating the limit of capital fluctuations, which is extremely friendly for novices and the confused. For example, when investors want to buy stocks of mature companies, pursuing stable returns, they can enter specific instructions in the agent and

wait for the result. The agent will automatically open the stock website and screen out the stocks that meet users' needs layer by layer, and then make a comprehensive and all-round decision after the investors have confirmed it's the one they are interested in. A wide range of investment options, flexible asset allocation, and professional and objective data provide the clients of a robo-advisor with more personalized choices, bringing more satisfaction.

Another apparent advantage of the robo-advisor is that it is open to the public and operates in real-time. The robo-advisor breaks with tradition by continuously collecting instant data, which helps to reduce the information gap. Traditional AI only refers to past events and trends, makes comparisons, and responds similarly. For events that have never occurred before, traditional AI has no way to respond.

However, the robo-advisor follows current news reports and constantly analyzes the globally changing situation. As it tracks open information, it can generate quick coping strategies to meet the challenges of emergencies. When the market changes significantly, the robo-advisor creates human-readable warnings and cautions with contextually appropriate explanatory content, which reminds investors in time so that they can prepare for the change ahead of schedule. According to Charles Schwab, a remarkable financial institution, the robo-advisor achieved a 28% retention rate and a 42% portfolio adjustment reduction, indicating that investors are willing to heed the advice and make correct decisions in the current situation. The result shows that through immediate adjustment, investors obtain significantly higher returns. First-hand information constantly protects investors from the mutated market. It also set aside enough time to deal with the sudden great changes.

Moreover, a robo-advisor has constant adaptability to studying. As the algorithm undergoes updates and iterations, the old algorithms will gradually be replaced.

Robo-advisor keeps pace with the times and optimizes computing power and models. So, facing personal finance management and robot consultation problems, robo-advisor transforms big data into personal value with the back of extraction of structured and unstructured data present to investors after integration [8].

Another cutting-edge application is the emotion-aware AI customer service that can be directly put into use. Based on users' questions and responses, generative AI can spontaneously detect users' emotions, then adjust the way of communication and solve the problem properly. It is applicable in a wide range of scenarios, like after-sales services, business handling, human relay work, and so on. The most valuable aspect lies in improving users satisfaction and convenience, due to the repeated confirmation of satisfaction level.

As emotion-aware AI customer service detects users' emotional fluctuations and explains and soothes emotions, while robo-advisors formulate rational investments for the clients, they address the emotional bias of financial decisions together. Additionally, the combination ensures a rational and logical decision, avoiding the extreme consequence caused by emotional overreaction. This can efficiently help customers to calm down in a short time, avoiding making extreme decisions by the higher-ups.

3.2 Efficient Operation and Product Development

The value of generative AI works in efficient operation embodied in cost reduction, efficiency lifting, and human endowment. As generative AI is highly skilled in operational content, analysis and decision making, it is highly suitable for improving efficiency throughout the entire process. It can play a significant role in intelligent customer service and AI&RPA (Robotic Process Automation) process automation, enhancing human-machine collaboration while strengthening the ability of scientific decision-making. At the same time, generative AI empowers the development thoroughly. After generating users profile, it automatically uses generative design and assists with algorithms and codes, creating multiplex mode products. Moreover, a trial-and-error type of experiment is put into practice to select the excellent ones from the large-scale test samples. It can break down the barriers of traditional industries and give rise to innovative products.

One typical skill of generative AI is drafting and reviewing contracts. Through database, industry templates, target requirement, etc., generative AI automatically and quickly generates a contract that meets the user's needs. Contrary to traditional human labor, generative AI comes with a strict review mechanism and is able to examine every flaw comprehensively. In addition, compared to making a contract manually, generative AI has achieved speed and accuracy, helping enterprises and individuals save significant time and cost. For companies and their staff, it effectively reduces the loopholes that are overlooked during manual review. Also, a standard contract can reduce potential loss and prevent problems related to non-compliance with regulations.

4. The Challenges and Methodology of Risk Prevention and Control in Finance

In the past, the detection of fraudulent activities mainly depended on manual work, comparing and screening out the abnormal data, which was particularly inefficient and prone to error. By contrast, generative AI leverages neural networks and anomaly detection models and studies the normal trading patterns to capture outliers in a large amount of data quickly and precisely. In recent years, there are fraudulent activities such as identity theft, payment fraud, account takeover by phishing, malware attacks, and social engineering. In the intricate transactions in the financial field, fraudulent activities have always been an awful thing for both investors and the market. Hence, how to prevent financial fraud is of utmost importance.

Baseline model and natural language processing (NLP) are two fundamental machine learning modes. NLP filters the data by the identity of named entities and keyword research and specializes in machine translation, question answering systems, language modelling, and text summarization [9]. The baseline model covers the definition standard, simplified implementation, and the boundaries definition. Generative AI makes a comparison between the actual transaction data and the baseline model and concludes whether this financial transaction has problems in a short time, which prominently saves people's energy and time.

Since robo-advisors receive most of their data from the public domain, which contains a great deal of complicated and redundant information, it's hard to ensure that every piece of information is correct. If the generative AI database is filled with false information, the content it generates will become horrible.

Nevertheless, due to the openness of the database, this approach still poses risks. Individuals with malicious intent may upload fake news and spurious transaction data online. Consequently, the generative system may be misled, and the baseline may deviate. Subsequently, the data they detect will no longer be accurate. Therefore, it is extremely important to promptly identify false information.

As mentioned in the data sources for robo-advisors, robo-advisor has delved into unstructured data such as social media, news reports, and voice communication. Generative AI generates simulated data through GAN and LLM, breaking through the predicament of traditional AI algorithms. So, it can identify and analyze underlying data in a highly efficient and accurate manner. For instance, generative AI can track current trends and respond to them appropriately, timely adjusting the investment proportions of individual financial products. It can also simulate extreme market scenarios and make corresponding adjustments to ensure that the risks for investors are minimized. Investors' risk appetites and pressure tests offer AI different scenarios to adapt to. By training on various scenarios and utilizing mathematics and algorithms, AI can evaluate the impact of extreme events and provide more appropriate advice to investors.

Despite having great potential, generative AI still faces some challenges that need to be addressed in the financial sector. The urgent problem that needs an immediate solution is how to provide AI with training data that does not violate privacy [10].

At the same time, regulatory institutions cannot precisely and comprehensively understand the AI operating model. Thus, whether there are any under-the-table dealings remains unknown. Therefore, the transparency of information and the effectiveness of supervision are the main factors influencing market confidence. More publicity should be made available to all.

To sum up, AI is promising to change the financial regulatory system by increasing the ability to identify risk, the capacity to enhance regulatory strategy and boost model and algorithm. However, to maximize its potential, regulators should address the problem of data privacy and accuracy and the ethical AI problems.

5. Conclusion

Generative AI has been evolving rapidly for many years. In the era of advancing science and technology, generative AI offers convenience and sound advice to people in the financial field based on databases, special algorithms, and models. AIGC and machine learning have laid the foundation for AI in the financial field, while robo-advisors and emotionally-aware AI have been integrated and are being widely used. Yet, the excessive growth of generative AI brings inevitable risks. Financial fraud always exists in the market, so how to address unexpected and sudden incidents becomes crucial. As the existing regulations cannot cover all aspects, explainable computer language models and necessary human intervention supervision prove effective.

In the future, generative AI will become more mature, cover a wider range of fields, and have closer interdisciplinary intersections with finance. In the long term, generative AI will continuously empower inclusive finance and promote the development of finance. Due to the rapid development of the Internet, the financial risks still need to be seriously, as algorithm black box and data forgery advent frequently, up-to-date supervision and constantly updated algorithms seem valuable. Governments and supervision institutions should play a leading role in the punishment and prevention of illegal use of generative AI, and build a periodic risk management system to ensure the security of online finance.

References

- [1] Oluwagbenro, M. B. (2024). Generative AI: Definition, concepts, applications, and future prospects. Authorea Preprints.
- [2] Udeh, E. O., Amajuoyi, P., Adeusi, K. B., & Scott, A. O. (2024). The role of big data in detecting and preventing financial fraud in digital transactions. *World Journal of Advanced Research and Reviews*, 22(2), 1746-1760.
- [3] Ridzuan, N. N., Masri, M., Anshari, M., Fitriyani, N. L., & Syafrudin, M. (2024). AI in the financial sector: The line between innovation, regulation and ethical responsibility. *Information*, 15(8), 432.
- [4] REMOLINA LEON, N. (2024). Generative AI in finance: risks and potential solutions. *Law, Ethics and Technology*, 1(1), 1.
- [5] Chen, B., Wu, Z., & Zhao, R. (2023). From fiction to fact: the growing role of generative AI in business and finance. *Journal of Chinese Economic and Business Studies*, 21(4), 471-496.
- [6] Lee, D. K. C., Guan, C., Yu, Y., & Ding, Q. (2024). A comprehensive review of generative AI in finance. *FinTech*, 3(3), 460-478.
- [7] Arenas-Parra, M., Rico-Pérez, H., & Quiroga-Garcia, R. (2024). The emerging field of Robo Advisor: A relational analysis. *Heliyon*, 10(16).
- [8] Anshari, M., Almunawar, M. N., & Masri, M. (2022). Digital twin: Financial technology's next frontier of robo-advisor. *Journal of risk and financial management*, 15(4), 163.
- [9] Sawicki, J., Ganzha, M., & Paprzycki, M. (2023). The state of the art of natural language processing—a systematic automated review of NLP literature using NLP techniques. *Data Intelligence*, 5(3), 707-749.
- [10] Akinrinola, O., Okoye, C. C., Ofodile, O. C., & Ugochukwu, C. E. (2024). Navigating and reviewing ethical dilemmas in AI development: Strategies for transparency, fairness, and accountability. *GSC Advanced Research and Reviews*, 18(3), 050-058.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

CSR Governance Framework for Internet Platform Enterprises

Rongjia Su and Quan Yuan*

Education and Training Institute, Sichuan University, China

**Corresponding author: Quan Yuan*

Abstract

The Internet platform economy is a new driving force for economic development, mobilizing entrepreneurial enthusiasm, activating social resources, and promoting integrated development across sectors. This article studies the CSR governance framework of Internet platform enterprises. From a multi-level governance perspective, it deconstructs governance priorities and approaches. This can not only fulfill social responsibility in every transaction of two-sided users of the platform, but also help platform enterprises expand the mission of “technology for good” from the field of public welfare to the overall operation of the platform.

Keywords

internet platform enterprises, ecological governance framework, CSR

1. Introduction

Internet platform enterprises refer to emerging organizations which build multi-sided, two-sided, and cross-industry trading platforms with the support of Internet technology to provide goods or services to market entities for a fee [1]. As the new business system nurtured by the Internet, Internet platform enterprises build infrastructure, connect supply-side enterprises and consumers, gather corresponding social and economic resources, and jointly create value with two-sided users [2]. As a special business ecosystem, interconnected platform enterprises do not provide products or services themselves, but to meet the needs of multi-sided and two-sided markets, break the constraints of time and space, promote the interaction between business users and consumer users, facilitate the transaction of products or services, and jointly create value. Platform enterprises can be roughly divided into trading platforms, information platforms, software platforms and social platforms according to the type of users.

Corporate social responsibility (CSR) is the voluntary activities of an enterprise in business operations, including concern for society and the environment, and interaction with stakeholders [3]. CSR Stakeholders not only include employees, investors, consumers, suppliers and competitors, but also society, government and the environment. The CSR dimensions, among which the most recognized are the four responsibilities summarized in the research of Carroll [4], economic responsibility, legal responsibility, ethical responsibility and philanthropic responsibility. If further categorized according to stakeholders, CSR can be divided into responsibilities to insiders, including responsibilities to employees, and to financial stakeholders,

responsibilities to external business partners, including responsibilities to consumers, suppliers, market responsibility, and business responsibility, and responsibilities to the public, including environmental responsibilities and philanthropic responsibilities.

Like traditional CSR, the CSR of the Internet platform also starts from economic responsibility and expands to legal, ethical and philanthropic responsibilities. However, because it involves two-sided users and the role of the Internet platform, the social responsibility of platform enterprises has a dual connotation. Internet platform enterprises should not only assume their own social responsibility, but also take social responsibility for the whole platform as the platform leader and coordinator. The CSR rules formulated and implemented by the platform have an impact on the platform enterprises themselves and two-sided users. Because the platform enterprises themselves do not provide products or services, but build and operate a platform, promote the interaction between business users and consumer users, and facilitate the transaction of products or services. Therefore, the target of platform enterprises to fulfill their social responsibilities has changed. Consumers and suppliers are no longer external stakeholders, but become internal stakeholders of platform enterprises, because their quantity and scale play a crucial role in the survival and development of platform enterprises.

Platform enterprises should not only actively fulfill their social responsibilities, but also guide the development of the platform business ecosystem into a sustainable ecosystem, and work with the government to regulate the CSR behavior of two-sided users of the platform. The network effect of the platform economy and the external characteristics distinguish the social responsibility of Internet platform enterprises from traditional enterprises. There are many ways for Internet platform enterprises to fulfill their social responsibilities. It can mobilize the platform's own existing resources to help solve some social problems. For instance, platform firms can fulfill employee responsibilities, including improving the working environment, strengthening social security, implementing equity incentive plans, and providing career development and training. They can also mobilize relevant enterprises, two-sided users and charitable organization to pay attention to social issues and build a socially responsible platform, such as environmental protection activities advocated by Ant Forest, and public fundraising organized by Waterdrop [5, 6]; or joining external CSR activities, such as cooperating with relevant departments to supervise and provide a safe network environment [7].

Although Internet platform enterprises also bear legal, environmental, economic responsibility and social responsibilities like traditional enterprises, the economic responsibility of platform enterprises involves more free and value-added Internet services for multi-market users. Ethical responsibilities include the construction of a socially beneficial intelligent society and environmental protection [1]. Legal responsibilities include combating pirated online resources. At the same time, platform enterprises also assume the responsibility to ensure security, which includes offline supervision, credit rating, and qualification verification [8].

It is also worth noting that when Internet platform enterprises are at different stages of development, they place different emphasis on the legal, environmental, economic and social responsibilities. In the early stage of development, platform enterprises will pay greater attention to economic responsibility, focus on expanding the size of two-sided users, optimize the efficiency of resource allocation, and reduce platform transaction costs, thereby promoting value creation, increasing corporate profits, and meeting the expectations of internal stakeholders. With the continuous development of the platform and the growth of the two-sided market, Internet platform enterprises will take the initiative to fulfill their legal and ethical responsibilities. Because young managers with an international vision and high ethical standards will take the initiative to abide by laws and regulations, pursue technological progress and innovation [1]. With the continuous growth of the platform, the number of users are relatively stable, the business ecosystem of the platform is increasingly mature, and the external stakeholders have higher expectations for Internet platform enterprises to fulfill social obligations. It is hoped that these enterprises can use social influence to address social issues and create social value.

2. Research Progress on CSR Governance on the Internet Platform

The operating environment of the Internet platform is different from that of traditional enterprises, the CSR governance model of the Internet platform gradually focuses on multi-stakeholder collaborative

governance. Scholars have built relevant governance paths from the perspectives of laws and regulations, internal monitoring and government supervision. In terms of laws and regulations, Gao [9] put forward the concept of third-party obligations under administrative law and proposed to combine third-party obligations with platform responsibility and platform governance. Regarding internal control, Chao, Lin, and Wang [10] used the leverage balance model to explain how platform enterprises fulfill their social responsibilities; among them, the support point of leverage is platform employees, and both sides of the leverage are two-sided users. When the status of two-sided users changes relatively, the leverage should be tilted, and platform enterprises need to change the way of management and elements of social responsibility to balance the leverage. Platform enterprises need to innovate the internal management model. Wang and Zhang [11] put forward the concept of greenhouse environment. The greenhouse environment is an intermediate layer between the internal management of the organization and the external environment. Platform enterprises need to design a small market within the large social market and establish a management mechanism for sellers to standardize their behaviour. These mechanisms are integrated to form a greenhouse management model. Platform enterprises should establish an effective greenhouse management model through the coordinated matching of rule settings, program settings, management arrangement and system integration. However, internal governance depends on the initiative and voluntary willingness of enterprises to implement social responsibility. Government supervision should consider the degree of intervention and the appropriateness of its intervention method [12]. At the level of government supervision, Liu and Xia [13] pointed out that when formulating regulatory policies, the government should adhere to the bottom-line thinking. When there are problems involving personal information leakage, potential safety risk, and market monopoly, it should solve specific problems, instead of completely denying the new business model. Li and Tao [14] and others mentioned the concept of fourth-party governance. They pointed out that, in addition to the characteristics of general enterprises, large-scale platform enterprises also have more public powers to provide public management services, and do not belong to the three types of legal persons defined by the General Principles of the Civil Law. Therefore, they refer to large-scale platform enterprises as fourth-party persons, and the various public powers they have are collectively referred to as fourth-party governance. They also pointed out that to ensure the implementation of fourth-party governance, the government needs to build an online supervision platform, and large platform enterprises must share their big data systems with the supervision platform in real time to gradually build a comprehensive big data supervision platform. The social responsibility of Internet platform enterprises not only involves economic issues at the market level, but also social issues related to the public sphere. It is difficult to achieve effective governance results by relying solely on individual enterprises, related industries or government agencies. Therefore, the concept of CSR governance on the platform needs to be transformed from one-sided governance in the individual context and multi-sided governance in the group context to collaborative governance in the Internet platform context. The multiple stakeholders of the Internet platform need to give full play to their respective advantages, cooperate with each other, and form a network of multi-stakeholder collaborative governance.

3. CSR of the Internet Platform and Diversified Governance Framework

The social responsibility of the platform enterprises can be managed from the three levels of individual, contextual and systemic levels, to establish the governance framework of individual, business ecosystem and social ecosystem, which can curb the internal factors that lead to the irresponsible behavior of individuals, and eliminate the external contextual and systemic factors that lead to the irresponsible behavior of enterprises [7]. When choosing a CSR governance framework, platform enterprises can not only adopt the individual CSR governance framework based on the platform's attributes, but also embed themselves within the social system [2].

3.1 Individual Level

At the Individual level, the targets of governance include platform enterprises at level I, two-sided users at level II and business ecosystem members at level III. The value orientation, behavioral logic, internal governance and ability to perform responsibilities of these individuals are all internal factors that lead to irresponsible social behaviour. Therefore, the key to the governance of CSR at the individual level lies in the guidance of values and the enhancement of CSR awareness, to spur the intrinsic consciousness and

motivation of individuals to fulfill their responsibilities, and to adopt a responsible behavior towards society [7].

At the individual level, when dealing with social and public problems, the individual resources of the enterprise are the main resources of governance. As far as the platform is concerned, the main focus is on its own basic functions. For the supporting members of the platform's business circle, it is necessary to maximize the support of stakeholders, such as shareholders, employees and other internal resources. CSR governance at the individual level should make full use of the rich resources and professional capabilities of the platform to actively participate in the governance of social problems in a direct or indirect way.

In the prevention and control of COVID-19, major Internet platform enterprises actively shoulder social responsibility and take the initiative to provide various support based on their own resources and capabilities to ensure public health. In addition, these Internet platform enterprises rely on their unique digital technology capabilities to promote social governance and innovation. The LinearFold algorithm developed by Baidu has greatly shortened the time of whole gene sequencing of the COVID-19 virus. Tencent, Alibaba, JD.com and other Internet platform enterprises have developed billions of computing power resources to provide computing support and big data analysis support for drug and vaccine research and development [15].

3.2 Contextual Level

At the contextual level, the activities of platform enterprises, two-sided users and other members constitute the platform business ecosystem. The operation model, competition practices, rules and the platform business ecosystem will affect the psychological and behavioral patterns of its members. Rules designed to merely pursue profit and vicious competition within the platform will weaken the collective CSR awareness in the platform's business ecosystem and cause irresponsible social behaviour of members of the business ecosystem. Therefore, the focus of CSR governance at the contextual level is to strengthen the collective social responsibility awareness of the business ecosystem and avoid the destruction of the value of the entire business ecosystem due to the irresponsible behavior of two-sided users. It is necessary to create a healthy, interactive, win-win business ecosystem.

Platform enterprises, which are in the core position of the platform business ecosystem, play the dual role of leadership and governance at the contextual level. They need to guide the dominant and supportive users in the main position of the platform to form behavioral patterns that are responsible for society. At the same time, it is necessary to go beyond individual-level resources and functions by acquiring, mobilizing, and integrating resources within the business ecosystem, thus realizing the dual governance framework of "platform individual + platform business ecosystem".

In the process of fighting the COVID-19 epidemic, major Internet platform enterprises exert influence on the behavior of users in the business ecosystem, give full play to the role played by platform enterprises, strictly supervise the medical products and services provided by sellers, and prevent any opportunistic behavior. Taobao, Pinduoduo, JD.com and other e-commerce platforms supervise and regulate the quality and price of epidemic prevention materials supplied by their platform sellers by establishing and implementing a negative list system and user evaluation system, and effectively prevent sellers from producing and selling poor-quality epidemic prevention materials, raising prices, and false advertising and other irresponsible behaviors to society [16].

3.3 Systemic Level

For the CSR governance at the systemic level, platform enterprises need to stand from the perspective of society, supervise the vicious competition between platforms through appropriate institutional guidance and arrangements, and create a benign competitive atmosphere; and optimize the allocation of social resources based on the business ecosystem, so that the business ecosystem can provide support for the social system, give full play to the role of social resource allocators under the social ecosystem, and encourage diverse participants in the social ecosystem to take part in the solution of economic, social and environmental problems, to realize the value creation of society and the environment.

During the epidemic, major e-commerce platforms rely on the professional ability of the business ecosystem and the advantages of rich resources to strengthen the willingness of ecosystem users and diverse

social actors to help prevent and control the epidemic, and encourage them to actively participate in solving social problems. E-commerce platforms such as Alipay and Meituan have included public welfare platforms, such as COVID-19 assistance platforms, donation information release platforms, and recipient platforms into their business systems, and use platform resources to call for diversified social actors to contribute to the supply of epidemic materials. Live streaming platforms, including Kuaishou and Douyin, make full use of the user resources of platform streamers to actively disseminate the knowledge of epidemic prevention and control to the public [16].

4. Conclusions

In addition to choosing the governance model from individual, contextual and systemic levels, platform enterprises can also choose different governance approaches according to different development stages. For example, in the pre-event stage, institutional governance can be selected to prevent irresponsible social behavior through the formulation and implementation of relevant systems. During the in-event stage, technology and organizational governance can play a role. Information technology can be used to establish a technological system with real-time monitoring to support the CSR governance; through organizational governance, the governance boundaries and priorities can be clarified. The post-event stage mainly relies on the institutional approaches. Internet platform enterprises need to design and implement rules and regulations applicable to platform transactions, which can not only play a preventive role, but also help deal with routine issues and emergencies, to ensure that members of the business circle have well-established channels for entry and exit, accountability mechanisms, reward and punishment systems.

References

- [1] Hu, Y., Hao, Y., & Chen, W. (2021). A Comparative Study on Corporate Social Responsibility between Internet Platform Enterprises and Traditional Manufacturing Enterprises: Based on the Background of Constructing a Dual Circulation New Development Pattern. *Journal of Chongqing University (Social Sciences Edition)*, 1-12.
- [2] Yang, Z. (2018). Corporate Social Responsibility of Platform Enterprises: Boundaries, Governance and Evaluation. *Economist*(5), 79-88.
- [3] Van Marrewijk, M. (2003). Concepts and definitions of CSR and corporate sustainability: between agency and communion. *Journal of Business Ethics*, 44(2-3), 95-105.
- [4] Carroll, A. B. (1991). The pyramid of corporate social responsibility: toward the moral management of organizational stakeholders. *Business Horizons*, 34(4), 39-48.
- [5] Zhu, X., & Li, M. (2020). Legitimacy and Content Analysis of CSR in E-commerce Platform Enterprises. *Social Science Research*, 28-36.
- [6] Liu, C., Guo, B., & Zhong, W. (2014). Collaborative Mechanism of Corporate Social Responsibility in China's E-commerce Retail Enterprises. *Economic Issues*, 60-63.
- [7] Xiao, H., & Li, P. (2019). Ecological Governance of Corporate Social Responsibility in Platform Enterprises. *Management World*, 35, 120-144.
- [8] Zhou, X. (2019). Safety Assurance Obligations and Legal Liability of E-commerce Platforms. *Academic Research*, 66-73.
- [9] Gao, Q. (2014). On Third-Party Obligations in Administrative Law. *East China University of Political Science and Law Journal*, 38-56.
- [10] Chao, G., Lin, D., & Wang, L. (2017). Corporate Social Responsibility Behavior Models of Platform Enterprises: A Case Study Based on Two-Sided Markets. *Journal of Management Case Studies*, 10, 70-86.
- [11] Wang, X., & Zhang, Q. (2016). Greenhouse Management Model of Platform-based E-commerce Enterprises: A Case Study of Alibaba Group. *China Industrial Economics*, 108-125.

- [12] Yang, Z., & Xu, Y. (2018). Governance of Corporate Social Responsibility under the Platform Economy. *Enterprise Economy*, 37, 78-86.
- [13] Liu, Y., & Xia, J. (2016). Research Trends in Sharing Economy Theory and Policy. *Economic Perspectives*, 116-125.
- [14] Li, G., & Tao, T. (2018). E-commerce Platform Ecosystem and Platform Governance Policy. *Management World*, 34, 104-109.
- [15] Qu, S., & Yang, D. (2020). Intelligent Response to Public Health Emergencies: Theoretical Review and Trend Analysis. *Reform*, 14-21.
- [16] Yang, Z., Yin, X., & Chen, J. (2020). Innovation in CSR Governance of Platform Enterprises under COVID-19. *Chinese Journal of Management*, 17.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

The Underlying Logic, Practical Challenges and Optimal Pathways for Urban Commercial Banks to Expand into and Empower the New Rural Collective Economy

Zhuoya Jiang*, Conghui Wu, Yaru Li and Liye Zhu

Guangxi Normal University, Guilin, Guangxi, China

**Corresponding author: Zhuoya Jiang*

Abstract

Can urban commercial banks effectively empower the development of the new rural collective economy by expanding into rural markets? This paper constructs a three-dimensional theoretical analytical framework of “cost–risk–efficiency”. Drawing on the typical case of Guilin Bank’s four-tier service network (city–county–township–village), and employing literature review and case study methods, it systematically analyses the micro-level mechanisms and practical challenges involved in urban commercial banks’ efforts to empower the rural collective economy. The research finds that, through mechanism innovations such as a four-tier service network, governance by rural elites, and industrial chain finance, Guilin Bank has demonstrated the theoretical feasibility of such empowerment; however, in practice, constraints such as the insufficient capacity of village collectives, high banking service costs, and an imperfect external institutional environment have hindered the full realisation of the empowerment effects. The marginal contribution of this paper lies in: adopting a problem-oriented perspective to reveal the practical dilemma that urban commercial banks face in empowering the rural collective economy—namely, that while theoretically feasible, implementation is hindered—and in providing targeted policy recommendations to promote the optimised development of this model.

Keywords

urban commercial banks, new rural collective economy, four-tier service network, cost–risk–efficiency, practical challenges

1. Introduction

The new rural collective economy serves as a vital vehicle for comprehensively advancing rural revitalisation and achieving common prosperity for farmers and rural communities. By the end of 2025, China’s rural population is projected to reach approximately 451 million [1], and rural areas possess a wealth of diverse resources, including land, ecological assets, cultural and tourism resources, and labour. Developing and strengthening the village-level collective economy plays an irreplaceable role in revitalising collective assets, boosting farmers’ incomes, and enhancing the effectiveness of rural governance [2]. In his report to the 20th National Congress of the Communist Party of China, General Secretary Xi Jinping emphasised: “We must consolidate and improve the basic rural operating system, develop the new rural collective economy, foster new

types of agricultural operating entities and socialised services, and promote moderate-scale agricultural operations” [3]. The 2025 Central Document No. 1 explicitly states that in developing a new type of rural collective economy, it is necessary to “innovate investment and financing mechanisms for rural revitalisation” and guide financial institutions to increase financial support for the rural collective economy [4].

However, a large number of villages in China currently face the dilemma of a weak collective economy and unstable operating income, and a shortage of funds remains one of the primary issues constraining the development of the value of village collective resources and the realisation of endogenous development. Although a series of inclusive finance and rural revitalisation support policies have been introduced at the national level, financial assistance to agriculture faces multiple challenges: rural collective economic organisations lack collateral, information asymmetry is pronounced, and there is a mismatch between the supply of financial products and actual demand, leading to financial institutions being “reluctant to lend, unwilling to lend, and unable to lend” [5]. How to overcome these challenges and ensure that finance truly benefits the rural collective economy is a key issue of common concern in the current context of rural revitalisation.

Finance is the core of the modern economy, playing a vital role in facilitating capital allocation and driving economic growth. The Notice on Effectively Carrying Out Financial Work for Agriculture, Rural Areas and Farmers in 2025, issued by the National Financial Supervisory Administration, explicitly states the need to “establish a rural financial service institution system characterised by clear tiers, complementary strengths and orderly competition” [6]. Against this backdrop, various financial institutions are accelerating their expansion into rural markets. Li Dan [7] points out that large banks, leveraging their scale of funds and network advantages, exert a “leading-goose effect” at the county level by increasing volume and reducing costs, whilst small and medium-sized rural financial institutions rely on their geographical advantages to cultivate the local market. However, existing research pays little attention to the expansion practices of city commercial banks as a specific category of institutions.

As local joint-stock commercial banks established through the restructuring of urban credit cooperatives and formed through equity participation by local governments and enterprises, urban commercial banks possess unique advantages such as proximity to the local community, short decision-making chains, timely access to information, and flexible service responses. In recent years, some urban commercial banks have begun to explore viable pathways for penetrating the rural market, with the “city–county–township–village” four-tier service network model established by Guilin Bank attracting particular attention. So, can the expansion of urban commercial banks into rural markets effectively empower the rural collective economy? What are the underlying mechanisms? What practical challenges do they face? And how might these be optimised?

It is worth reflecting on the fact that feasibility in theory does not necessarily equate to feasibility in practice. The validity of theoretical logic relies on a series of preconditions, whereas the practical level often faces multiple obstacles. Consequently, this study not only examines the micro-level mechanisms through which city commercial banks empower the rural collective economy, but also focuses on analysing the practical challenges encountered in the transition from theory to practice, with particular emphasis on dissecting the complex duality of “mechanism feasibility” coexisting with “practical difficulties”.

Compared with existing research, the marginal contributions of this paper are as follows: firstly, it focuses on city commercial banks as a specific entity, avoiding conflation with large commercial banks and rural commercial banks, thereby clarifying the scope of the study; secondly, adopting a problem-oriented perspective to reveal the dual nature of urban commercial banks’ empowerment of the rural collective economy—namely, that it is “theoretically feasible but hindered in practice”—rather than merely providing a simple feasibility argument; and thirdly, drawing on practical case studies from Guilin Bank to conduct an in-depth analysis of the empowerment mechanisms and practical challenges, and to propose targeted policy recommendations.

2. Literature Review and Analytical Framework

2.1 Literature Review

2.1.1 Research Progress on the Empowerment of Rural Collective Economies by Financial Institutions

Academic research on the role of financial institutions in empowering the rural collective economy has accumulated a substantial body of work. Overall, whilst the extension of financial services to rural areas has achieved notable results, it continues to face challenges on multiple fronts. In response, scholars have analysed the situation from various perspectives. Firstly, on the demand side, rural collective economic organisations face the pressing issues of “difficult, expensive and slow access to finance”. Gao [8] points out that there is controversy regarding the legal interpretation of the guarantee capacity of rural collective economic organisations as special legal entities, making it difficult for financial institutions to regard them as eligible borrowers; furthermore, He [9] emphasise in their research that the problem of information asymmetry is pronounced, and the fragmentation of credit information regarding business entities exacerbates the lending concerns of financial institutions. Secondly, on the supply side, there is a structural mismatch between the supply of financial services and the demand of the rural collective economy. Traditional financial institutions still primarily offer small-scale unsecured loans, making it difficult to meet the scaled and long-term funding needs of new types of agricultural business entities. Wang et al. [10] therefore point out that the homogenisation of credit products is a significant issue; Furthermore, Liu et al. [11] suggest that financial institutions tend to overestimate rural financial risks—such as low project profitability and high risk—leading to a lack of initiative in serving the rural collective economy and strict controls on credit volumes.

Regarding the causes of the aforementioned challenges, scholars have conducted multifaceted analyses. Among them, Xiao et al. [12] emphasise that the lag in rural collective property rights reform, coupled with unclear ownership definitions and an underdeveloped transfer market, makes it difficult to convert assets into effective collateral; Guo et al. [13], in their empirical research, pointed out that in the practice of financial institutions extending their reach into rural areas, there is a structural contradiction between commercial operations and the outcomes of agricultural support, making it difficult to balance social and economic benefits; thirdly, Kang et al. [14] argued that the lack of sound risk-sharing and compensation mechanisms, coupled with the lagging development of rural credit reporting systems, results in high credit acquisition costs and difficulties in risk assessment for financial institutions; Fourthly, multi-stakeholder collaboration mechanisms are underdeveloped; communication channels between government, banking and business entities involved in financial outreach have not yet become institutionalised, resulting in a lack of synergy during the empowerment process [15].

Gao et al. [18] emphasise the need to strengthen financial safeguards for the rural collective economy and improve the system for financing through the mortgaging of collective assets; Zhu and Sun [19], drawing on practical case studies from multiple regions, propose innovative models for financial support of the new rural collective economy, including loans secured by equity pledges of collective assets, industrial chain finance, collective development and joint credit facilities; furthermore, Peng et al. [20], from the perspective of policy optimisation, propose that a mechanism for fiscal and financial coordination should be established to leverage financial capital towards the rural collective economy sector through fiscal interest subsidies and risk compensation.

2.1.2 Research Progress on the Expansion of Urban Commercial Banks into Rural Areas

In recent years, some scholars have begun to focus on the practices of urban commercial banks in expanding into rural markets. He and Wang [19] utilised county-level clustering data on Guilin Bank’s rural inclusive finance service points to examine whether urban commercial banks’ expansion into rural areas can balance agricultural support performance with operational performance. The study indicates that Guilin Bank’s four-tier service network mechanism in rural areas can simultaneously enhance both agricultural support performance and operational performance, whilst the participation of rural elites can significantly strengthen the mechanism’s impact on agricultural support; mechanism analysis reveals that alleviating information asymmetry, effectively controlling agricultural support costs, and establishing sound risk-sharing mechanisms are key pathways to improving both types of performance.

Yin [20] analysed the situation from the perspective of the expansion of inclusive finance into rural areas, pointing out that the expansion of city commercial banks into rural areas is a strategic choice made against the backdrop of saturated competition in urban markets, with their core advantages lying in technological investment, product innovation and risk control capabilities; at the same time, they face fierce competition

from the Postal Savings Bank of China, the Agricultural Bank of China and local rural commercial banks, necessitating the adoption of a service strategy based on specialisation and differentiation.

Existing research provides valuable insights for understanding financial empowerment of the rural collective economy; however, much of this research tends to lump large banks, rural commercial banks and city commercial banks together under the umbrella term “financial institutions” for analysis, failing to give sufficient attention to the distinctive strengths and specific challenges faced by city commercial banks in serving the rural collective economy. Mechanism analysis and challenge analysis are treated as separate entities. Existing research either focuses on demonstrating the feasibility of financial empowerment or remains confined to a macro-level description of challenges, lacking a research perspective that organically integrates mechanism analysis with an analysis of challenges. There is a lack of in-depth micro-level examination based on typical case studies. Macro-level policy recommendations struggle to guide specific practical operations; it is necessary to conduct an in-depth analysis of the operational logic and practical obstacles of empowerment mechanisms through typical case studies.

This paper aims to address the aforementioned research shortcomings by adopting a problem-oriented approach. Using Guilin Bank’s outreach mechanism as a case study, it systematically analyses the intrinsic mechanisms and practical challenges faced by city commercial banks in extending services to the rural collective economy, and proposes corresponding recommendations for optimisation.

2.2 Theoretical Framework

This paper will utilise a three-dimensional analytical framework centred on “cost–risk–efficiency”, incorporating key concepts from the theories of information asymmetry, transaction costs and financial symbiosis, to explore the underlying logic and practical pathways for city commercial banks to extend their reach and empower the rural collective economy. By addressing the gap between theoretical feasibility and practical implementation, the paper aims to establish a comprehensive theoretical analytical framework that both underpins the validity of the framework and clarifies the core analytical perspective of this study.

2.2.1 The “Cost–Risk–Efficiency” Three-Dimensional Analysis Framework

Within the analytical framework, costs encompass operational costs, customer acquisition costs, data collection costs and management costs. To expand their reach into rural areas, city commercial banks must reduce these costs through innovative business models, such as adopting a light-asset service point model, relying on rural elites to reduce data collection costs, and utilising digital technology to lower the cost of processing individual transactions. Risks include credit risk, market risk and operational risk. The rural collective economy faces issues such as an underdeveloped credit system, weak operational stability and a lack of collateral, necessitating the establishment of a multi-tiered risk mitigation mechanism. Efficiency encompasses service delivery efficiency, loan approval efficiency and capital allocation efficiency. It is necessary to improve the alignment between financial services and the needs of the rural collective economy by shortening service distances, optimising approval processes and precisely matching demand.

2.2.2 Theoretical Foundation

The theory of information asymmetry suggests that the imbalance of information between lenders and borrowers in credit markets is prone to trigger adverse selection and moral hazard, leading financial institutions to implement credit rationing [21]. Furthermore, issues such as financial irregularities and a lack of creditworthiness within rural collective economic organisations create significant information barriers with financial institutions, thereby exacerbating this predicament. Transaction cost theory, meanwhile, demonstrates that transaction costs throughout the entire process—including information search, negotiation and contract formation—determine the feasibility and commercial sustainability of market transactions [22]. The characteristics of rural collective economy credit—namely small individual loan sizes, decentralised services and long cycles—keep transaction costs under standardised credit models persistently high, constituting the core reason for financial institutions’ low willingness to supply such credit. The theory of financial symbiosis proposes that financial institutions and entities in the real economy can form interdependent, mutually beneficial relationships characterised by synergy, comprising three key elements: symbiotic units, models and environments. This provides a theoretical framework for understanding the relationship between city commercial banks and rural collective economies [23]. Whilst the two are

potentially compatible symbiotic units, the establishment of their symbiotic relationship requires the support of a sound external institutional environment. Financial empowerment is not a one-way process of supply and demand, but rather requires both parties to form a symbiotic structure of collaborative development.

The expansion of city commercial banks' services into the rural collective economy essentially involves the dynamic balancing and optimisation of the relationship between cost, risk and efficiency. These three factors are intertwined and determine the feasibility and sustainability of such service expansion. The inherent characteristics of credit for the rural collective economy drive up the end-to-end transaction costs for city commercial banks, whilst information asymmetry further exacerbates credit risk. The combination of high costs and high risks can trap city commercial banks in a dilemma of inefficiency, thereby dampening their enthusiasm for expanding into rural areas.

Conversely, if city commercial banks can reduce various costs through business model optimisation, improved risk control and the application of financial technology, whilst overcoming information barriers and credit allocation challenges, they can achieve cost reductions whilst maintaining controllable risks. This, in turn, enhances the efficiency of both service delivery and capital utilisation, creating a virtuous cycle of low costs, low risks and high efficiency. This not only assists the rural collective economy in securing credit support but also enables the city commercial banks to achieve their own differentiated development. This also validates the explanatory power of the three major theories regarding the practical constraints on urban commercial banks' expansion into rural areas.

The logical chain through which city commercial banks empower rural collective economies is: geographical advantage→ institutional innovation→ triple performance→ sustainable development. The smooth operation of this chain, however, requires numerous preconditions. This paper takes the outreach mechanism of Guilin Bank as its analytical case study, exploring both the successful experiences of its institutional innovation and the practical challenges encountered in implementation. It aims to provide a reference that combines theoretical depth with practical value, thereby enabling city commercial banks to better empower rural collective economies.

3. A Logical Analysis of How Urban Commercial Banks Expand into Rural Areas to Empower the Rural Collective Economy

3.1 Guilin Bank's Four-Tier Service Network: Organisational

The four-tier service network established by Guilin Bank—spanning city, county, township and village levels—serves as the core vehicle for its expansion into rural areas. This service network demonstrates significant institutional innovation in terms of cost control.

Firstly, a light-asset approach reduces the costs of branch construction and operation. Guilin Bank has broken away from the traditional heavy-asset branch model of conventional banks. Village-level inclusive finance service points are prioritised for establishment in the homes of villagers with good reputations within the village, utilising their own premises to provide services. This eliminates the need for additional leasing or construction of business premises, thereby substantially reducing the costs of market development and fixed operational costs in rural areas. These service points are equipped to assist farmers, enabling the centralised handling of basic transactions such as small-value deposits and withdrawals, social security payments and account enquiries. This effectively diverts basic business from branch outlets to village-level service points, significantly reducing the transaction costs per transaction.

Secondly, a tiered coordination system reduces management and coordination costs. Guilin Bank has established a clear mechanism for the division of responsibilities within the four-tier network: the municipal branch is responsible for product development, policy formulation and the coordination of risk control; county-level branches are responsible for project coordination, credit approval and local management; township micro-branches are responsible for identifying village collective needs, pre-loan due diligence and post-loan management; and village-level service points are responsible for information collection, reporting needs and providing basic services. This has formed a closed-loop collaborative system characterised by “city-level coordination, county-level implementation, township-level outreach and village-level frontline presence”.

Thirdly, optimising localised funding has reduced funding costs. Relying on its four-tier service network, Guilin Bank has extensively attracted low-cost deposits from county and rural areas. According to Guilin Bank's 2025 interim report, as of the end of June 2025, the Bank's deposit balance stood at 374.95 billion yuan; During the reporting period, the Bank continued to deepen the penetration of its county-level lending operations, with the outstanding balance of county-level loans reaching 118.342 billion yuan and the outstanding balance of agriculture-related loans standing at 59.914 billion yuan, representing year-on-year increases of 7.78% and 10.27% respectively. By extending its four-tier service network into rural areas, the Bank has effectively optimised its liability structure, creating scope for it to provide low-interest credit products to rural collective economies [24].

By establishing a four-tier service network spanning “city–county–township–village”, Guilin Bank has extended financial services to the village level. The adoption of a light-asset deployment model has significantly reduced the operational costs of financial services, enabling the bank to provide credit support to rural collectives at lower interest rates; simultaneously, a localised liability strategy has further optimised the structure of funding sources, continuously injecting stable financial resources into rural economic development. This organisational innovation has not only effectively improved the accessibility of financial services in rural areas but has also reshaped the rural financial ecosystem, laying a solid foundation for the modernisation of the rural collective economy.

3.2 Governance by rural elites: Mitigating Information Asymmetry

The active involvement of village elites is a distinctive feature of the Guilin Bank model. By carefully selecting respected and well-informed local figures to serve as service point managers, the bank has established a risk identification and prevention mechanism that integrates “soft” and “hard” information, thereby effectively alleviating information asymmetry between the village collective and individual farmers. In theory, this mechanism offers an innovative approach to overcoming information barriers in rural finance.

Firstly, the selection of station managers and the incentive and accountability mechanisms ensure the authenticity of information. Guilin Bank prioritises selecting village “Two Committees” officials, local entrepreneurs, and veteran Party members as station managers for its village-level service points. These individuals are familiar with the actual circumstances of every household and every village collective project, enabling them to accurately grasp the village collective's operational status, project prospects, creditworthiness and other “soft information” that is difficult to quantify. At the same time, Guilin Bank has established an incentive mechanism for station managers, directly linking their earnings to service quality and credit risk. This not only motivates station managers to proactively gather information and match needs, but also effectively prevents issues such as information falsification and moral hazard. By the end of June 2025, Guilin Bank had established nearly 6,000 rural inclusive finance service points across Guangxi, with station managers carefully selected from respected and influential figures within the villages, thereby forming an information-gathering network covering the entire rural area of Guangxi.

Secondly, the integration of soft and hard data enables end-to-end risk control. During the pre-loan phase, station managers can conduct preliminary due diligence on village collective projects through their personal networks, combining soft data—such as the collective's reputation, governance standards and project feasibility—with hard data held by the bank, such as credit records and collective asset data, to form a comprehensive credit profile. During the loan approval phase, station managers can participate throughout the evaluation of village collective projects, providing localised insights to support the bank's decision-making; during the post-loan phase, station managers can monitor operational changes and fund utilisation in village collective projects in real time, identifying potential risks immediately and reporting them to the bank.

Thirdly, the social trust network lowers the barrier to financial services for village collectives. Rural elites act as a bond of trust between the bank and village collectives, breaking down the trust barrier between rural residents and financial institutions, and effectively resolving the issues of village collectives “daring not to borrow” and “not knowing how to borrow”. Station managers can also use their daily services to educate village collective leaders on financial knowledge and explain credit products, assisting the village collectives in completing loan applications and preparing documentation.

By selecting rural elites familiar with local conditions and public sentiment to serve as service point managers, Guilin Bank has effectively alleviated the information asymmetry between the bank and rural

collectives. Leveraging their geographical advantages and the trust of the local community, these managers can accurately assess the operational status and creditworthiness of village collectives, thereby reducing the bank's credit risk and transaction costs. Building on this foundation, the bank utilises a mechanism integrating soft and hard data to gain a more comprehensive and accurate understanding of the actual situation of rural collective economic organisations, thereby providing more targeted and precise financial support. The establishment of a network of trust based on personal connections has further enhanced financial accessibility for village collectives, effectively stimulating the endogenous development momentum of the rural collective economy and promoting the deep integration of rural governance with modern finance.

3.3 Supply Chain Finance: Improving Efficiency and Sharing Risk

Guilin Bank has moved beyond the traditional, single-dimensional model of credit-based empowerment by deeply embedding financial services within distinctive rural industrial chains. It has established a service system comprising “credit support + production and marketing empowerment + risk-sharing”, which not only ensures a precise match between financial services and the needs of village collective industries but also achieves risk diversification through the coordination of multiple stakeholders across the industrial chain. Whilst this mechanism demonstrates the theoretical potential for comprehensive empowerment, it faces significant barriers to replication in practice.

Firstly, bespoke industrial chain products enable the precise allocation of funds. Guilin Bank has focused on distinctive agricultural industrial chains such as the Baise mango, Guilin mandarin orange and Liuzhou snail noodle ingredient sectors in Guangxi. By conducting in-depth research into the operational characteristics and funding requirements of each link in the chain, the bank has tailored exclusive credit products for local industrial chain projects. Taking the Baise mango industry chain as an example, Guilin Bank has launched the “Funan Loan – Mango Loan” product, covering the entire chain from mango cultivation to storage, preservation, processing and sales. This breaks away from the traditional reliance on collateral, instead granting credit facilities based on the industry chain operational data and order contracts held by village collectives. As of the end of June 2025, the outstanding credit facilities granted by Guilin Bank's Baise Branch to clients in the mango industry chain reached 62.76 million yuan, benefiting 714 growers, with unsecured loans accounting for as much as 86% of the total.

Secondly, end-to-end services enhance the efficiency of collective economic development. Guilin Bank's service scope has expanded beyond mere capital provision to encompass the entire industrial chain, extending to both upstream and downstream sectors. Leveraging “Finance+” agricultural support platforms such as the “Tiantian Kaixin Group” community group-buying platform and rural revitalisation livestreaming studios, the bank has established a “Bank + Leading Enterprises + Farmers” production-to-sales linkage model to assist village collectives in broadening their agricultural product sales channels.

Thirdly, multi-party collaboration across the industrial chain has established a risk-sharing mechanism. Through a multi-party cooperation model involving “leading enterprises + banks + farmers”, Guilin Bank has linked entities across the industrial chain, thereby achieving the dispersion and shared bearing of credit risks. Leveraging the credit endorsement and order commitments of leading enterprises, this model disperses the credit risk of individual loans to village collectives across the entire industrial chain; by collaborating with government-backed financing guarantee institutions and agricultural insurance companies, and introducing tools such as credit enhancement through guarantees and insurance hedging, a multi-tiered risk mitigation system comprising “credit + guarantee + insurance” has been established.

Based on the above analysis, Guilin Bank's four-tier service network, rural elite governance, and industrial chain finance mechanisms provide a viable pathway for city commercial banks to empower rural collective economies in terms of cost control, information mitigation, efficiency enhancement and risk sharing. However, institutional innovation does not necessarily equate to successful implementation. As demonstrated by Guilin Bank's experience, the effectiveness of these three mechanisms is highly dependent on a series of prerequisites: the light-asset model requires village collectives to possess basic financial needs; the elite governance mechanism requires the sustained participation and effective incentivisation of rural elites; and industrial chain finance requires a sound industrial foundation and the leadership of leading

enterprises. When these prerequisites are not met, the efficacy of the mechanisms will be significantly diminished.

4. The Practical Challenges Faced by City Commercial Banks in Empowering the Rural Collective Economy

The preceding section analysed, based on the case of Guilin Bank, how a four-tier service network achieves cost control, how governance by rural elites alleviates information asymmetry, and how industrial chain finance enhances efficiency and shares risks. Theoretical analysis indicates that city commercial banks possess inherent advantages in serving the rural collective economy, and that the institutional design demonstrates a certain degree of feasibility. However, there remain multiple practical obstacles between theoretical logic and practical implementation. This chapter will, in light of these practical challenges, systematically analyse the key constraints facing city commercial banks in empowering the rural collective economy from the three dimensions of cost, risk and efficiency.

4.1 Cost-related Challenges: The Operational Pressures on the Service Network Continue to Mount

Although the four-tier service network adopted by city commercial banks is innovative in terms of cost control, it faces multiple pressures in practice. Firstly, the effectiveness of the light-asset model is highly dependent on the activity levels and transaction volumes at service points. If the collective economic development in the villages covered by a service point lags behind and financial demand is insufficient, it will be difficult to cover the operating costs of the service point, leaving the bank in the awkward position of having built facilities that are not being utilised. Furthermore, the effectiveness of tiered coordination relies on efficient collaboration across all levels. As the number of service points increases and coverage expands, information transmission between tiers may become distorted or delayed, and coordination costs may rise rather than fall. Striking a balance between scale expansion and coordination efficiency is a practical challenge facing Guilin Bank. The cost advantage of county-level deposits may narrow as market competition intensifies. Should other financial institutions simultaneously step up their efforts to attract deposits in the county-level market, the competitive cost of such deposits will rise, potentially eroding Guilin Bank's advantage in funding costs.

4.2 Governance by Rural Elites: Uncertainties Surrounding the Implementation of the Model

Elite-led governance in rural areas has proven highly effective in mitigating information asymmetry; however, the stability and replicability of such mechanisms in practice face numerous challenges. The willingness of rural elites to participate in rural financial services is influenced by multiple factors: if the incentives are insufficient to cover their time costs, their participation may wane; if the incentive mechanisms are poorly designed, they may give rise to conflicts of interest between elites and farmers. The effectiveness of integrating soft and hard information depends on the comprehensiveness and accuracy of data collection. However, information on village collective operations is often poorly standardised and lacks timeliness; relying solely on the individual capabilities of station managers makes it difficult to guarantee the completeness and timeliness of information. Furthermore, familiar networks may be subject to issues such as "favour-based loans" or "connection-based loans", and soft information may be distorted by subjective emotions. Moreover, establishing a trust network based on personal connections requires a long period of accumulation and has a limited scope of effectiveness. In newly entered market areas, the cost of building trust networks is high; for business expansion across villages or townships, the effectiveness of familiar-based trust may diminish. This governance model, which relies heavily on "people", will face bottlenecks during replication and scaling.

4.3 The Dual Dilemma of Efficiency and Risk: High Barriers to Replicating Supply Chain Finance

The development of bespoke industrial chain products requires substantial investment in preliminary research and product design, resulting in high R&D costs; furthermore, product solutions are highly

dependent on the operational characteristics of specific industrial chains, making replication across different chains challenging. For regions with weak industrial foundations and incomplete industrial chains, the replicability of this model remains questionable. The extension of full-chain services requires banks to possess cross-sectoral resource integration capabilities, including the operation of e-commerce platforms, live-streaming sales and cold-chain logistics. For city commercial banks that primarily focus on traditional deposit and loan business, the transition to becoming comprehensive service providers faces multiple challenges, including talent reserves, technical capabilities and organisational structure. The effective operation of the industrial chain risk-sharing mechanism relies heavily on the willingness of leading enterprises to cooperate and the coverage capacity of guarantee institutions. However, leading enterprises are not universally present; for village collectives lacking the driving force of such enterprises, this mechanism is difficult to apply. Furthermore, the business focus of financing guarantee institutions remains on small and micro-enterprises and farmers, resulting in limited coverage of village collective economic organisations.

5. Advice

Based on the analysis of the aforementioned challenges, this paper argues that, in order to better empower the development of new rural collective economies, urban commercial banks need to make concerted efforts across three dimensions—model optimisation, capacity building and institutional improvement—to formulate a systematic policy framework.

5.1 Optimising the Cost Dimension: Promoting the Sustainable Development of the Service Network

In response to the pressure of service network operating costs, it is recommended that efforts be focused on the following areas: firstly, establishing a mechanism for evaluating the efficiency of service points. Implement categorised management based on indicators such as service point transaction volumes and the number of rural collectives covered, and consolidate and optimise underperforming service points to avoid a situation where “construction is prioritised over operations”. Second, explore diversified revenue models for service points. Whilst ensuring the provision of basic financial services, expand into value-added services such as insurance agency, agricultural product sales agency and financial literacy promotion to enhance the self-sustaining capacity of service points. Third, deepen bank-government cooperation to secure policy support. By signing strategic cooperation agreements with local governments, service point development should be incorporated into the overall rural revitalisation plan to secure policy subsidies for premises and operations.

5.2 Strengthening the Risk Dimension: Improving Incentive and Constraint Mechanisms for Elite Governance

To address the issue of stability in rural elite governance, the following improvements are recommended. Firstly, establish a certification system for station managers. Conduct regular professional training for station managers and grant certification to those who pass the assessment, thereby enhancing both their professional capabilities and their sense of belonging and honour. Second, optimise the incentive mechanism for station managers. Building upon the existing model of “basic subsidy plus business commission”, explore ways to link station managers’ earnings to the long-term development of the villages they serve. Third, establish standardised protocols for the collection of soft data. Develop a credit assessment template for village collectives, converting the soft data collected by station managers into quantifiable, traceable standardised data, thereby reducing the risks of interference from personal relationships and information distortion. Fourth, establish a rotation and succession mechanism for station managers. This will prevent service points from becoming overly reliant on a single station manager, cultivate a pool of reserve station managers, and ensure the sustainability of the elite governance model.

5.3 Enhancing Efficiency: Lowering the Barriers to Replicating Industrial Chain Finance

To address the challenges in replicating industrial chain finance, the following breakthroughs are recommended: Firstly, build a repository of industrial chain finance service templates. Systematically summarise the core elements of successful cases such as the Mango Loan and Sando Orange Loan to create

replicable product templates, thereby reducing product development costs for new industrial chains. Second, establish a cross-sector resource integration platform. Forge strategic partnerships with e-commerce platforms, logistics firms and agritech companies to avoid each bank having to build their own infrastructure from scratch. Third, explore a “One Village, One Product” differentiated service strategy. For regions lacking a complete industrial chain, guide city commercial banks to focus on a single link or a single product, using a “point-to-area” approach to gradually extend the service chain, thereby reducing the pressure of transitioning into comprehensive service providers.

6. Conclusion

In summary, this study analyses the logic and challenges underlying city commercial banks’ efforts to empower the rural collective economy within the “cost–risk–efficiency” framework. Theoretically, it fills a gap in research on the differentiation of city commercial banks, whilst in practice it provides practical case references for financial support to agriculture. Compared to existing research, which tends to conflate different types of financial institutions and separate the analysis of mechanisms from that of challenges, this study focuses on city commercial banks as the primary subject and achieves an organic integration of the two. However, it has yet to resolve the core challenge of replicating this model in underdeveloped regions with weak industrial foundations. The relevant recommendations for optimisation provide practical insights for the expansion of city commercial banks into rural areas and the improvement of rural financial systems. Future research could further explore pathways for integrating digital finance with the expansion models of city commercial banks, thereby promoting the deep integration and coordinated development of financial services and the rural collective economy.

References

- [1] National Bureau of Statistics. Wang Pingping: China’s total population will stand at 1, 404.89 million in 2025; high-quality population development continues to advance. Available from: https://www.stats.gov.cn/sj/sjzd/202601/t20260119_1962338.html [Accessed 25th March 2026].
- [2] Du Jiao. The Current State, Challenges and Strategies for the Development of Rural Collective Economies in China [J]. *Agricultural Economy*, 2021, (01):41-43.
- [3] Xinhua News Agency. Xi Jinping: Hold High the Great Banner of Socialism with Chinese Characteristics and Strive Unitedly to Build a Modern Socialist Country in All Respects — Report to the 20th National Congress of the Communist Party of China. Available from: https://www.gov.cn/xinwen/2022-10/25/content_5721685.htm [Accessed 25th March 2026].
- [4] Xinhua News Agency. Opinions of the Central Committee of the Communist Party of China and the State Council on Further Deepening Rural Reform and Steadily Advancing the Comprehensive Revitalisation of Rural Areas. Available from: https://www.moa.gov.cn/10443/ztl/2025yhwj/2025nzyyhj/202502/t20250224_6470428.htm [Accessed 25th March 2026].
- [5] Li Jianzhong, Li Shuang, Yuan Shaobiao. Challenges in Providing Financial Support for Rural Revitalisation [J]. *China Finance*, 2019, (24):102.
- [6] National Financial Regulatory Administration. The National Financial Regulatory Administration has issued the ‘Notice on Effectively Carrying Out Financial Work Related to Agriculture, Rural Areas and Farmers in 2025’. Available from: https://www.gov.cn/lianbo/bumen/202504/content_7019984.htm [Accessed 25th March 2026].
- [7] Li Dan. Harnessing the ‘leading-geese effect’ of major state-owned banks: An interview with the head of the Inclusive Finance Division at the China Construction Bank [J]. *China Financialist*, 2020, (07):48-50.
- [8] Gao Shengping. On the Procedural Controls and Legal Consequences of Guarantees Provided by Rural Collective Economic Organisations: Focusing on the Improvement of the Draft Law of the People’s

- Republic of China on Rural Collective Economic Organisations [J]. *Academia Bimestris*, 2024, (03):59-70+214.
- [9] Research Group of People's Bank of China, Guyuan Branch, He Wenhui. The Financing Predicament and Path Choice of Collective Economic Organizations: Taking 484 Rural Collective Economic Organizations in Ningxia as an Example [J]. *Financial Development Review*, 2025, No.183(03):82-95.
- [10] Wang Xiuhua, Wei Nianying. Rural Collective Economy Financing: Dilemma and Countermeasures [J]. *Issues in Agricultural Economy*, 2024, (03):72-88.
- [11] Liu Ming, Han Jingjing, Ge Weiwei. Economic Opportunity, Relationship-based Financing and Credit Allocation in Underdeveloped Rural Areas in Western China: An Analysis based on the Investigative Data from 1138 Rural Families in Shaanxi and Qinghai [J]. *Journal of Shaanxi Normal University(Philosophy and Social Sciences Edition)*, 2012, 41(04):102-112.
- [12] Xiao Weidong, Sheng Sheng. Equity-backed loans secured against rural collective assets: operational mechanisms, practical models and optimisation strategies [J]. *Dongyue Tribune*, 2025, 46(06):107-116.
- [13] Guo Yan, Han Qingxiao. Profitability, Agricultural Support Services and Risk Control: A Theoretical Analysis and Empirical Examination of Scale Adjustment in Rural Commercial Banks [J]. *Journal of Financial Research*, 2019, (04):130-148.
- [14] Kang Yaming, Lin Ziguang, Liu Ran. Research on the Whole Village Credit Service Model in China: Taking Beijing, Heilongjiang, Zhejiang and Guangdong Provinces and Municipalities as Examples [J]. *Management for Economy in Agricultural Scientific Research*, 2024, (02):44-48.
- [15] Chen Guangli. Practical Challenges and Optimal Approaches to Financial Support for the High-Quality Development of Rural Industries: A Case Study of Guangxi [J]. *China Collective Economy*, 2026, (07):5-8.
- [16] Gao Ming, Song Hongyuan, Jiang Fan. Path Selection and Policy Optimization of Finance Supporting the Development of Rural Collective Economy [J]. *Rural Finance Research*, 2022, (02):3-12.
- [17] Zhu Ruibo, Sun Chunlin. New Rural Collective Economic Innovative Development and Financial Support: Practical Cases, Problems and Suggestions [J]. *Southwest Finance*, 2023, (12):54-66.
- [18] Peng Peng, Wu Minhui, Zhang Longyao. The Theoretical Logic and Implementation Mechanism of Fiscal and Financial Synergy Support for the Development of Village-Level Collective Economy: Based on the Case Study of "Xingcun Easy-Loan" in Xinghua, Jiangsu Province [J]. *Issues in Agricultural Economy*, 2024, (08):59-72.
- [19] He Li, Wang Xiuhua. Can Urban Commercial Banks' Rural Outreach Balance the Performance of Supporting Agriculture and the Bank's Business Performance? Evidences from Guilin Bank's Four-Level Service Network [J]. *Contemporary Finance & Economics*, 2025, (06):72-84.
- [20] Yin Ziyang. A Study on the Adaptability of Human Resources and the Enhancement of Economic Efficiency in Urban Commercial Banks against the Background of the Expansion of Inclusive Finance [J]. *Shangxun*, 2025, (21):81-83.
- [21] Joseph E. Stiglitz and Andrew Weiss. Credit Rationing in Markets with Imperfect Information[J]. *The American Economic Review*, 1981, 71(3):393-410.
- [22] Williamson, O. E. *The economic institutions of capitalism: Firms, markets, relational contracting*. New York: Free Press;1985
- [23] Yuan Chunqing. Symbiosis Theory and Its Application to Small-Scale Economies (Part 1) [J]. *Reform*, 1998, (02):100-104.
- [24] China Money Network. Guilin Bank Co., Ltd. 2025 Interim Report. Available from: <https://www.shibor.org/chinese/cwbg/20250829/3186923.html> [Accessed 25th March 2026].

Funding

This paper was supported by supported by Guangxi Normal University's National Training Program of Innovation and Entrepreneurship for Undergraduates. Project A Study on the Mechanisms and Pathways for Enhancing the Role of City Commercial Banks' Decentralised Operations in Empowering the Development of New Rural Collective Economies (No. X2025106020090).

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

The Impact of 5G Factory Construction on Corporate Financial Efficiency— An Empirical Analysis Based on Manufacturing Listed Companies on the Shanghai and Shenzhen A-Share Markets

Ziyao Song*

Business School, Soochow University, Suzhou 215000, China

**Corresponding author: Ziyao Song.*

Abstract

In the era of digital economy, 5G factories, as the core carrier of intelligent manufacturing, represent an important manifestation of the transformation and upgrading of the manufacturing industry. This paper takes the manufacturing listed companies in Shanghai and Shenzhen stock markets from 2020 to 2024 as the sample. Using the “5G Factory List” released by the Ministry of Industry and Information Technology as the experimental group, a multi-period difference-in-differences model is employed to empirically examine the impact of 5G factory construction on the financial efficiency of enterprises. The research found that the construction of 5G factories significantly enhanced the total factor productivity of manufacturing enterprises. This conclusion remained robust after undergoing Parallel Trend Tests, Propensity Score Matching–Difference-in-Differences Model, and Placebo Tests. The Mechanism Test shows that although individual production, inventory or innovation indicators did not show significant changes in the short term, the 5G factory improved overall efficiency through systematic resource allocation optimization. This study provides microeconomic evidence for understanding the empowerment of the real economy by digital technologies.

Keywords

5G factory, total factor productivity, difference-in-differences method, digital transformation

1. Introduction

The academic community has conducted numerous theoretical and empirical studies on the relationship between digital transformation and enterprise efficiency. The “task model” proposed by Acemoglu and Restrepo in the early days laid the theoretical foundation for the assertion that automation technologies enhance productivity [1]. After the evolution of digital technology, digital transformation has been regarded as the core driving force for high-quality economic development. Shostak et al. pointed out that it not only changes the organizational operations, but also enhances overall efficiency through technological innovation [2].

At the micro-enterprise level, the financial efficiency empowerment effect of digital transformation has been well verified: Moolkham et al. have confirmed that it can increase enterprise value by enhancing financial efficiency and effectiveness. Mavlutova et al. believe that it can boost operational efficiency. Ajmal et al. have

discovered that emerging digital technologies can improve enterprise profitability and optimize inventory efficiency [3-5]. Hossain pointed out that the Internet of Things can enhance the efficiency of enterprise investment [6].

2. Literature Review and Research Hypotheses

In terms of the influencing mechanism, the generation of innovation and the allocation of resources are the key paths: Yang Peng et al. found that digital technology can enhance knowledge acquisition and collaborative innovation capabilities, and improve the innovation efficiency of enterprises. Li Xuan believes that it can alleviate financing constraints and promote innovative output. Liu et al. confirmed that the digital transformation of finance can solve information asymmetry, reduce costs, and improve financial performance [7-9].

Regarding China's manufacturing industry, the studies conducted by Cao Xiaojing et al., Wu Xianfu et al., and Zhao Chen all demonstrate that digital transformation can significantly boost the total factor productivity of manufacturing enterprises [10-12]. The synergy of digital finance cannot be ignored either, as it provides macro-environmental support for related construction [13, 14]. However, digital transformation also has potential negative impacts. Bai Peiwen and Yu Li found that it might intensify market competition and reduce the price markup of enterprises [15]. It can be seen that the impact of 5G factory construction on the financial efficiency of enterprises is a complex system optimization process, and it is not a simple linear relationship.

In the era of digital economy, it is widely recognized that automation and intelligent technologies have reshaped the production function. Acemoglu and Restrepo's "task model" has laid the theoretical foundation for the enhancement of production efficiency through automation technologies [1]. 5G factories, as an advanced form of industrial internet, restructure the manufacturing process based on the characteristics of 5G technology. In theory, this can reduce marginal costs and optimize resource allocation. Based on this, an assumption is proposed.

H1: The construction of 5G factories significantly enhances the financial efficiency of manufacturing listed companies.

Based on the relevant research on how digital transformation enhances enterprise efficiency, this paper identifies three core paths through which 5G factories affect financial efficiency. On the path of improving production efficiency, digital transformation serves as a booster for operational efficiency improvement, and the application of industrial internet can further drive the growth of total factor productivity in the manufacturing industry [4, 10, 11]. The 5G factory achieves intelligent transformation of the entire production process, realizes equipment interconnection and process optimization, and enhances output while reducing losses. On the path of supply chain optimization, emerging digital technologies can optimize inventory efficiency, and the Internet of Things can enhance supply chain collaboration and investment efficiency [5, 6]. 5G factories break through the data barriers of each link in the supply chain, enabling real-time perception of demand and dynamic optimization of inventory. In the path of R&D innovation promotion, digital transformation enhances enterprise efficiency through technological innovation. Digital technology can also improve innovation efficiency, alleviate financing constraints, and promote innovation [2, 7, 8]. The 5G factory is not only an enterprise's digital and intelligent innovation practice but also provides technical and data support for R&D. Based on this, Hypothesis 2 is proposed.

H2: The 5G factory achieves a synergistic improvement in the financial efficiency of manufacturing enterprises through three paths: enhancing production efficiency, optimizing the supply chain, and promoting R&D innovation.

3. Research Design

This article selects the manufacturing companies listed on the Shanghai and Shenzhen stock exchanges from 2020 to 2024 as the initial sample. The 5G factory data is derived from the "2023 5G Factory List" and "2024 5G Factory List" released by the Ministry of Industry and Information Technology. The financial data is sourced from the CSMAR and Wind databases. After eliminating samples of the ST type, financial type and

those with missing values, and conducting 1% and 99% tail trimming on the continuous variables, a total of 13,663 observations were finally obtained.

The core dependent variable of this article is total factor productivity (fe_tfp), and the explanatory variables are the difference-in-differences terms (did). The specific variable definitions are shown in Table 1.

Table 1: Variable Definition Table

Variable type	Variable name	Symbol	Explanation
The dependent variable	Total Factor Productivity	fe_tfp	The TFP calculated using the LP method reflects the efficiency of input-output.
Explanatory variable	5G Factory Effect	did	Treatment×Period
Control variables	Enterprise size	$size$	The natural logarithm of the total assets at the end of the period
	Asset-liability ratio	lev	Total liabilities at the end of the period / Total assets at the end of the period
	Shareholding Concentration	$top1$	The shareholding ratio of the largest shareholder
	Research and development investment	rd	Research and development expenditure / operating revenue
	Growth potential	$growth$	Growth rate of operating income
	Independence	$Indep$	Number of independent directors / Total number of the board of directors
	Enterprise age	age	Observation year - Year of enterprise establishment + 1
	Production efficiency	$product$	Per capita operating income
Mechanism variable	Supply chain efficiency	$inventory$	Inventory turnover rate
	Innovation capability	$patent$	The logarithm of the number of annual patent applications increased by 1

Data source: Compiled based on CSMAR and Wind databases.

This paper constructs the following multi-period difference-in-differences model:

$$FE_{i,t} = \alpha_0 + \alpha_1 \cdot DID_{i,t} + \sum \alpha_k \cdot Controls_{i,t} + \mu_i + \lambda_t + \varepsilon_{i,t} \quad (1)$$

Among them, fe_tfp represents the dependent variable, α_1 is the core concern coefficient, and μ_i and λ_t are the individual and year fixed effects respectively.

4. Analysis of Empirical Results

4.1 Descriptive Statistics

Table 2 presents the descriptive statistics of the main variables. The mean of total factor productivity is 8.959, with a standard deviation of 1.012, indicating significant differences in efficiency among manufacturing enterprises. The mean of the core explanatory variables is 0.0126, suggesting that approximately 1.26% of the observations in the sample are in the state after the implementation of the 5G factory policy. This is consistent with the reality that 5G factories in China are still in the pilot and promotion stage.

Table 2: Descriptive Statistics of Key Variables

Variable	Sample size	Mean value	Standard deviation	Minimum value	Maximum value
fe_tfp	13,663	8.959	1.012	6.955	11.850
did	13,663	0.0126	0.111	0.000	1.000
$size$	13,663	22.24	1.153	20.19	25.84
lev	13,663	0.393	0.188	0.061	0.866
age	13,663	23.48	5.784	10.00	40.00
rd	13,663	6.171	5.578	0.160	34.34
$growth$	13,663	0.169	0.453	-0.624	2.515
$top1$	13,663	31.60	13.85	8.359	69.92
$indep$	13,663	0.393	0.077	0.250	0.600

Note: The data has been rounded to the 1st and 99th decimal places.

4.2 Baseline Regression Results

Table 3 presents the benchmark regression results. Column (1) shows that, after controlling for the individual fixed effects of enterprises, the fixed effects of years, and relevant control variables, the coefficient of the core explanatory variable “did” is 0.040, and it is statistically significant at the 1% level ($t = 2.61$). This means that, compared to the control group, the construction of 5G factories has significantly increased the total factor productivity of manufacturing enterprises by approximately 4%, verifying hypothesis H1, which states that the construction of 5G factories significantly enhances the financial efficiency of manufacturing listed companies. In terms of controlling variables, the size of the enterprise significantly promotes TFP, while the debt ratio and the intensity of R&D investment have a negative impact. This might be related to the fact that the current investments have not been promptly converted into outputs.

Table 3: Baseline Regression Results

Variable	(1)fe tfp
did	0.040***
	(2.61)
size	0.588***
	(31.18)
lev	-0.120**
	(-2.40)
rd	-0.050***
	(-19.89)
growth	-0.015**
	(-1.97)
top1	0.001
	(1.06)
indep	-0.015
	(-0.50)
cons	-3.788***
	(-8.96)
Sample size	13,427
R-squared	0.976
Individual fixed effect	YES
Fixed effect of year	YES

Note: The values in parentheses are t-values. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

4.3 Robustness test

4.3.1 Parallel trend test, PSM-DID test and placebo test

In order to verify the underlying assumptions of the DID model, this paper examines the dynamic effects before and after the implementation of the policy. The results shown in Table 4 indicate that before the implementation of the policy (Pre_3), the coefficient was significantly negative (-0.052), suggesting that the experimental group had lower efficiency before the establishment of the factory and there was a certain “catch-up effect”. However, in the year of policy implementation (Current) and subsequent years (Post_1), the coefficient turned positive. Although due to the short sample period, the statistical significance was not fully conclusive, the overall trend indicates that the construction of 5G factories has reversed the efficiency disadvantage, conforming to the dynamic characteristics of the parallel trend assumption, and further verifying the rationality of the DID model set in this paper.

Table 4: Results of Parallel Trend Test

Variable	Coefficient	t-value	$P > t $
pre_3	-0.052**	-2.05	0.040
pre_2	-0.036*	-1.94	0.052
current	0.014	0.98	0.328
post_1	0.003	0.10	0.921
post_2	(omitted)	-	-

Considering that the declaration of the 5G factory list is not completely random, there might be a “self-selection bias”, meaning that enterprises with larger scale and better management are more likely to be selected as 5G factories. To address this endogenous issue, this paper employs a combination of propensity score matching and difference-in-differences methods for the test.

This study examines whether an enterprise is a 5G factory (treatment) as the dependent variable, and selects control variables such as enterprise size and debt ratio (lev) as covariates for Logit regression. The 1:1 nearest neighbor matching method is used to find featureally similar control groups for the experimental group.

The matching result table 5 shows that after matching, the vast majority of samples are within the common support domain, while only a very small number of samples fall outside this range. This indicates that the matching effect is good, and the feature differences between the experimental group and the control group have been effectively eliminated. The average treatment effect (ATT) for the treatment group was 0.036. This value is very close to the coefficient of 0.040 in the baseline regression, indicating that after eliminating the sample selection bias, the positive impact of 5G factory construction on enterprise efficiency remains robust. The results of the benchmark regression are not solely driven by the inherent qualities of the sample, but rather are indeed attributed to the policy effects of the construction of 5G factories.

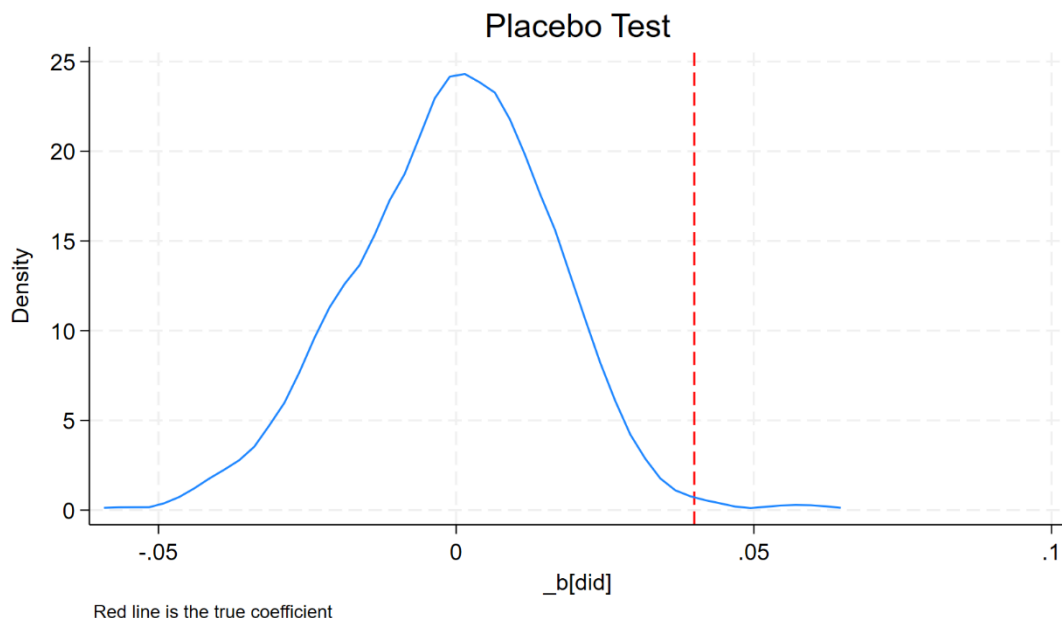
Table 5: PSM-DID Matching Results

Variable	Unmatched	ATT	Difference	S.E.	T-stat
fe_tfp	9.967	9.960	0.036	0.075	0.48

In order to further eliminate the interference of unobservable omitted variables on the research conclusions and ensure that the baseline regression results are not caused by random factors, this paper employs the Random Permutation Test for placebo verification.

Specifically, this paper maintains the control variables and fixed effects unchanged, randomly reassigns the allocation of the core explanatory variable “did” among the sample enterprises, constructs a false experimental group and a policy time interaction term, and then conducts regression estimation again. This process was repeated for 500 times to obtain 500 sets of false regression coefficient estimates. The results of the placebo test are shown in Figure 1.

Figure 1: Kernel density distribution map of the placebo test



As can be seen from Figure 1, the false regression coefficients generated through 500 random simulations (the blue curve) exhibit a standard normal distribution centered at 0. This indicates that, after randomly altering the processing status, the average impact effect of 5G factory construction on the financial efficiency of enterprises approaches zero.

The red dotted line in the figure represents the true estimated coefficient (0.040) of the baseline regression in this article. It can be clearly observed that this true coefficient is located at the far right tail of the kernel density distribution graph, completely separate from the randomly generated coefficient distribution area. According to the statistical results of the Monte Carlo permutation test, in 500 random samplings, only 2 simulations had coefficients greater than the actual observed values. The two-sided test's P-value was 0.008. This indicates that the probability of obtaining the current significant positive result under random conditions is less than 1%.

In conclusion, the core conclusion of this paper - “The construction of 5G factories enhances the financial efficiency of enterprises” - is not the result of accidental factors or unobservable omitted variables. The research conclusion is highly robust.

4.3.2 Change the Dependent Variable

Replace TFP with asset net profit ratio (ROA), total asset turnover ratio (TAT), and cost and expense profit margin (CPM). The results shown in Table 6 indicate that the 5G factory had no significant impact on these individual financial indicators. This result contrasts with the significant increase in TFP observed in the baseline regression, revealing the structural characteristics of the enabling effect of 5G factories. The significant increase in total factor productivity indicates that 5G technology has indeed optimized the production process, reduced redundant losses, and enhanced “technical efficiency”. However, no significant improvement was observed in the single-factor financial indicators (particularly ROA and TAT). The possible reasons for this are as follows: The construction of 5G factories is a capital-intensive investment. During the initial stage of construction, enterprises need to purchase expensive intelligent equipment and base stations (resulting in an increase in the denominator of total assets), and the conversion of technological dividends into book net profits has a certain lag. Therefore, in the short term, 5G factories mainly represent a “qualitative” leap in the technical efficiency of the production and manufacturing process, rather than an immediate and significant increase in the accounting financial return rate. This further indicates that using TFP (Total Factor Productivity) as the core indicator can more accurately capture the real efficiency contribution of digital transformation.

Table 6: Regression Results of Replaced Dependent Variable

Variable	(1)fe roa	(2)fe tat	(3)fe cpm
did	0.001	-0.003	0.009
	(0.34)	(-0.31)	(0.80)
Controls	YES	YES	YES
FE	YES	YES	YES
N	13,427	13,427	13,427
R-squared	0.744	0.905	0.792

4.4 Mechanism Verification

To verify the research hypothesis H2, this paper examined three paths of influence: production efficiency (product), supply chain efficiency (inventory), and innovation capability (patent). The results in Table 7 showed that the DID coefficients corresponding to the three mechanism variables were all not significant. Therefore, the research hypothesis H2 was not supported by the empirical results. The reason for this outcome lies in the fact that the construction of 5G factories is a systematic digital transformation that covers the entire production and operation chain of the enterprise. From the commissioning of intelligent equipment, the adaptation of personnel skills to the full-scale reconfiguration and implementation usually requires a 1-3 year integration period. The list of 5G factories in this sample was released in 2023, and the observation period only extends to 2024. The observation window after policy implementation is relatively short. In the short term, the efficiency improvement effect of a single link is difficult to manifest. At the same time, indicators such as patent applications and inventory turnover rate have strong lagging characteristics and cannot fully capture the enabling effect of 5G technology in the short term. This result also complements and corroborates the baseline regression, indicating that the improvement in enterprise efficiency by 5G factories is not achieved through an immediate burst in a single link, but is ultimately reflected as an improvement in total factor productivity through the optimization of systematic resource allocation throughout the entire chain. Currently, enterprises

are still in the transitional period of digital transformation, and the improvement of a single indicator has a significant time lag.

Table 7: Results of Mechanism Test Regression Analysis

Variable	(1) product	(2) inventory	(3) patent
did	-42111.2	-0.115	0.007
	(-0.25)	(-0.75)	(0.06)
Controls	YES	YES	YES
FE	YES	YES	YES
N	13,427	13,427	13,427

Note: product refers to per capita revenue, inventory refers to inventory turnover rate, and patent refers to the number of patent applications.

5. Conclusion

This paper takes the manufacturing companies listed on the Shanghai and Shenzhen stock exchanges from 2020 to 2024 as the sample. Based on the “2023 5G Factory List” of the Ministry of Industry and Information Technology, a quasi-natural experiment was constructed. Using the multi-period difference-in-differences model for empirical testing, it was found that the construction of 5G factories can significantly enhance the total factor productivity of manufacturing enterprises. This conclusion still holds after undergoing a series of robustness tests; At the same time, this efficiency improvement exhibits significant structural and systematic characteristics. In the short term, it mainly manifests as the improvement of technical efficiency at the production end. It has not yet fully translated into an increase in book financial performance, and it is not achieved through immediate breakthroughs in single links such as production, supply chain, and innovation. Instead, it stems from the systematic optimization of resource allocation throughout the entire chain. It is recommended that enterprises maintain their strategic determination for digital transformation and attach importance to the long-term systematic benefits of 5G factory construction.

This study still has certain limitations: Firstly, the observation window of the sample is relatively short, covering only the data from 1 to 2 years after the implementation of the policy, which makes it difficult to fully capture the long-term dynamic effects of the construction of 5G factories; Secondly, the core explanatory variables only use the dummy variable indicating whether a company is included in the list, failing to distinguish the construction scale and application depth of 5G factories; Thirdly, the sample only covers A-share manufacturing listed companies and does not include small and micro enterprises, so the generalizability of the conclusions needs to be expanded; Fourthly, the mechanism research only examines three basic paths and fails to further decompose the specific channels for systematic resource allocation optimization.

In response to the aforementioned limitations, future research can proceed in the following ways: First, extend the sample observation period to clarify the long-term change patterns and profit conversion cycles of the 5G factory empowerment effect; second, refine the multi-dimensional measurement indicators for 5G factory construction, and conduct heterogeneity analyses for different enterprises and industries; third, expand the sample coverage to include small and micro enterprises to complete the research on all entities; fourth, deepen the examination of the mechanism of action, and simultaneously expand the research perspective to further explore the impact of 5G factory construction on enterprise risks, green transformation, and industrial chain competitiveness, and more comprehensively evaluate the comprehensive value of 5G factory construction.

References

- [1] Acemoglu, D., & Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6), 1488-1542.
- [2] Shostak, L., Goi, V., Timchenko, O., Yastrubetska, L., & Derhaliuk, M. (2023). The impact of digital transformation on the economy: Technological innovation and efficiency. *Economic Affairs*, 68(4), 2081-2093.

- [3] Moolkham, M., Petchchedchoo, P., Kumsuprom, S., & Panmanee, P. (2025). Digital transformation, financial efficiency, financial effectiveness and firm values. *KKBS Journal of Business Administration and Accountancy*, 9(3).
- [4] Mavlutova, I., Spilbergs, A., Verdenhofs, A., Natrins, A., Arefjevs, I., & Volkova, T. (2023). Digital transformation as a driver of the financial sector sustainable development: An impact on financial inclusion and operational efficiency. *Sustainability*, 15(1), 207.
- [5] Ajmal, K., Wang, C. X., Suresh, N. C., & Vedantam, A. (2026). Impact of emerging digital technologies on firms' financial performance, inventory efficiency, and greenhouse gas emissions: An event study. *Sustainability*, 18(1), 1600.
- [6] Hossain, M. A., & Hossain, M. N. (2023). The influence of IoT and digital technologies on financial risk monitoring and investment efficiency in global supply chains. *American Journal of Interdisciplinary Studies*, 4(2), 91-125.
- [7] Yang Peng, Yin Zhifeng, Sun Baowen. (2024). Application of Digital Technologies in Enterprises and Enhancement of Innovation Efficiency. *Foreign Economics and Management*, 46(11).
- [8] Li Xuan. (2024). Research on the Impact of Digital Transformation on Enterprise Innovation Output: With Financing Constraints as the Mediating Variable. *Technology Entrepreneurship Monthly*, 37(12).
- [9] Liu, J., Zhou, K., Zhang, Y., & Tang, F. (2023). The effect of financial digital transformation on financial performance: The intermediary effect of information symmetry and operating costs. *Sustainability*, 15(6), 5059.
- [10] Zhao Xiaojing, Zhou Wenhui, Liu Yaying. (2024). Industrial Internet and Total Factor Productivity of Manufacturing Enterprises. *Economic Management*, (7), 5-22.
- [11] Wu Xianfu, Zhong Peng. (2024). The Impact of Industrial Internet on the Total Factor Productivity of Manufacturing Enterprises. *Consumption Economy*, (3), 20-31.
- [12] Zhao Chen. (2025). Research on the Impact of Digital Transformation on the Total Factor Productivity of Enterprises. *Accounting Friends*, (23).
- [13] He, Q., & Xue, Y. (2023). Research on the influence of digital finance on the economic efficiency of energy industry in the background of artificial intelligence. *Scientific Reports*, 13(1), 14984.
- [14] Syed, A. A., Özen, E., & Kamal, M. A. (2023). Do digital financial services influence banking stability and efficiency: An ARDL analysis of a developed and a developing economy. In *The New Digital Era: Digitalisation, Emerging Risks and Opportunities* (Vol. 109, pp. 13-30). Emerald Publishing Limited.
- [15] Bai Peiwen, Yu Li. (2021). Digital Economic Development and Enterprise Price Markup: Theoretical Mechanism and Empirical Facts. *China Industrial Economics*, (11), 59-77.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

The Impact of New Quality Productive Forces on Firm Performance and Its Mechanisms

Ruixuan Zhang*

Sun Wah International Business School, Liaoning University, Shenyang 110136, China

**Corresponding author: Ruixuan Zhang.*

Abstract

This study takes Chinese A-share listed firms in the Shanghai and Shenzhen stock markets from 2015 to 2024 as the research sample, constructs a panel data model, and empirically examines the impact of new quality productive forces on firms' short-term operating performance and long-term value performance. It further analyzes the underlying mechanisms, robustness, and heterogeneity characteristics. The results show that, first, new quality productive forces have a significantly positive effect on firm performance, and this conclusion remains robust after a series of robustness tests, including the use of instrumental variable approaches and alternative measurements of core variables. Second, mechanism analysis indicates that R&D investment and digital capability serve as important transmission channels through which new quality productive forces affect firm performance. Specifically, new quality productive forces improve firm performance by promoting increased R&D investment and enhancing digital technological innovation capability. Third, heterogeneity analysis reveals that the performance-enhancing effect of new quality productive forces is particularly pronounced in non-heavily polluting industries. At the regional level, a clear gradient pattern is observed: the promoting effect is stronger for firms located in regions with higher levels of economic development than those in less developed regions, and it is more significant for firms in the eastern region than in the central and western regions. This study extends the literature on the microeconomic consequences of new quality productive forces to a certain extent and provides new empirical evidence for understanding the mechanisms through which technological innovation and the optimized allocation of production factors contribute to firm value creation. Meanwhile, it offers practical implications for guiding firms to accurately seize development opportunities associated with new quality productive forces and for assisting governments in implementing differentiated policies.

Keywords

new quality productive forces, firm performance, digital capability, R&D investment

1. Introduction

At present, a new round of technological revolution and industrial transformation is deepening. Emerging technologies represented by artificial intelligence, big data, and quantum information continue to achieve breakthroughs, driving the development of productive forces into a new stage characterized by new quality productive forces as the core driving force. In September 2023, General Secretary Xi Jinping first proposed the concept of "new quality productive forces" during an inspection in Heilongjiang Province, emphasizing

the need to “vigorously develop strategic emerging industries such as new energy, new materials, advanced manufacturing, and electronic information, accelerate the development of future industries, and expedite the formation of new quality productive forces” [1]. In January 2024, during the eleventh collective study session of the Political Bureau of the CPC Central Committee, he further pointed out: “In simple terms, new quality productive forces are those in which innovation plays a leading role, break away from the traditional models of economic growth and productivity development, exhibit characteristics of high technology, high efficiency, and high quality, and conform to the new development philosophy as an advanced form of productive forces” [2]. Subsequently, the 2024 Government Work Report also explicitly proposed “accelerating the development of new quality productive forces” and “promoting green and low-carbon development” [3]. As a significant leap in the form of productive forces, new quality productive forces are increasingly permeating various fields of economic and social development and have become a key driver for enterprise transformation, upgrading, and high-quality development. Against this background, a systematic analysis of the impact of new quality productive forces on firm performance is of great significance for deepening relevant theoretical research and providing empirical evidence for high-quality firm development.

Following the introduction of the concept of new quality productive forces, the academic community has begun to examine its economic consequences at the micro firm level. Existing studies generally argue that new quality productive forces represent an advanced form of productivity characterized by technological breakthroughs, innovative allocation of production factors, and industrial structure upgrading. By enhancing resource allocation efficiency and innovation capability, they can improve firm performance. Ha jinhua et al. [4] find that firms can cultivate new quality productive forces through network embeddedness, which facilitates technological breakthroughs and the optimal allocation of production factors, thereby significantly improving firm performance. From the perspective of financial performance, new quality productive forces exert a significant impact on firm value and exhibit a clear threshold effect under different levels of profitability [5]. Some studies analyze this issue from the perspective of firm development capability, suggesting that new quality productive forces can enhance firms’ growth capacity and operational stability by improving resource utilization efficiency and optimizing the structure of resource allocation, thereby indirectly improving operating performance [6].

Overall, the existing literature has laid an important foundation for understanding the economic effects of new quality productive forces. Current studies mainly proceed along three lines. First, they explore the conceptual connotation and theoretical foundations of new quality productive forces, providing systematic definitions of their formation logic and core characteristics. Second, they analyze the mechanisms through which new quality productive forces operate, particularly from the perspectives of resource allocation efficiency, technological innovation capability, and industrial structure upgrading. Third, they evaluate the economic impact of new quality productive forces using outcome variables such as firm value, resilience, and growth. However, current research remains relatively fragmented in terms of perspectives and analytical frameworks, often focusing on single dimensions or specific contexts. The comprehensive effects and multidimensional economic consequences of new quality productive forces at the firm level still require further in-depth investigation. Moreover, some studies focus primarily on macroeconomic or industry-level analyses, and systematic micro-level empirical evidence at the firm level remains insufficient.

In view of these research gaps, this study selects Chinese A-share listed firms in the Shanghai and Shenzhen stock markets as the research sample and systematically examines the impact of new quality productive forces on firm performance and its underlying mechanisms. Compared with existing studies, the marginal contributions of this paper are threefold. First, in terms of research content, this study investigates the impact of new quality productive forces on firm performance at the micro level and further explores the mechanisms from the perspectives of R&D investment and digital capability, thereby extending the research on the economic consequences of new quality productive forces. Second, in terms of methodology, this study employs panel data of Chinese A-share listed firms over multiple years for empirical analysis and strengthens the reliability of the conclusions through various robustness and endogeneity tests, providing new empirical evidence for understanding the microeconomic effects of new quality productive forces. Third, in terms of research perspective, this study incorporates both short-term operating performance and long-term value performance into a unified analytical framework, offering a more comprehensive examination of the role of new quality productive forces in firm development and enriching the literature on the relationship between new quality productive forces and firm performance.

2. Theoretical Analysis and Research Hypotheses

2.1 New Quality Productive Forces and Firm Performance

New quality productive forces are innovation-driven and represent an advanced form of productivity that breaks away from traditional growth models. They are characterized by high technology, high efficiency, and high quality, and are generated by revolutionary technological breakthroughs, innovative allocation of production factors, and deep industrial transformation and upgrading [7, 8]. For firms, the enhancement of new quality productive forces is mainly reflected in the improvement of labor quality, the upgrading of means of production, and the expansion of objects of labor. Through the optimized combination of these elements, overall efficiency is improved, as evidenced by the continuous growth of total factor productivity. This transformation may promote both short-term profitability and long-term value through multiple channels. The following analysis examines the mechanisms through which new quality productive forces enhance firm performance from the perspectives of stakeholder theory, strategic management theory, and dynamic capability theory.

From the perspective of stakeholder theory, new quality productive forces help improve the relationships between firms and multiple stakeholders, thereby enhancing firm performance. The development of new quality productive forces requires firms to place greater emphasis on attracting and cultivating high-quality talent, which increases employees' human capital value and labor productivity [9]. At the same time, the green attributes of new quality productive forces promote green technological innovation, reduce environmental compliance costs, and expand green consumption markets, thereby generating higher sales revenue. The increased information transparency and strengthened responsibility fulfillment brought about by new quality productive forces help enhance the trust of stakeholders such as investors and consumers, which is reflected in improved firm market value [10]. The development of new quality productive forces can directly improve resource utilization efficiency and enhance collaborative effects with stakeholders such as employees and society, thereby promoting firm performance [6]. In addition, new quality productive forces encourage firms to strengthen safety production and quality management, safeguard consumer rights, reduce operational risks, and provide stable expectations for stakeholders, thereby laying a solid foundation for sustained improvements in firm performance [11].

From the perspective of strategic management theory, new quality productive forces themselves constitute strategic resources for firms to build differentiated competitive advantages, and their cultivation and accumulation directly shape firms' competitive positions. By integrating new types of production factors—such as digital technologies, green technologies, and high-quality human capital—into the traditional factor system, firms can occupy unique niches in industrial competition and directly obtain higher pricing power and profit margins. This strategic uniqueness, formed through the integration of new production factors, can directly enhance firms' market value and profitability [12]. The development of digital new quality productive forces also directly strengthens the resilience of industrial chains, as reflected in improved stability and sustained growth capacity in response to market fluctuations [13]. Strategic accumulation of production factors contributes to short-term sales growth through new product development and market expansion, while also ensuring long-term excess returns by establishing technological barriers.

From the perspective of dynamic capability theory, new quality productive forces themselves embody firms' higher-order capabilities to integrate, build, and reconfigure resources in response to environmental changes. The enhancement of dynamic capabilities enables firms to more accurately perceive technological changes and more efficiently allocate resources, and these advantages are directly translated into improved operational efficiency and financial performance. Firms' learning capabilities developed within competitive and cooperative networks, as an important component of new quality productive forces, can enhance their ability to adapt to environmental changes, thereby improving both financial and non-financial performance [14]. Strengthening internal learning orientation and external network relationships enables firms to better leverage new quality productive forces in dynamic environments, thereby directly promoting performance growth [15].

Based on the above analysis, this study proposes the following hypothesis:

H1: New quality productive forces have a significantly positive impact on firm performance. Specifically, the higher the level of new quality productive forces, the better the firm's short-term operating performance and long-term value performance.

2.2 New Quality Productive Forces, Digital Capability, and Firm Performance

Digital capability refers to a firm's ability to integrate internal and external resources and achieve value creation. It is typically reflected in the firm's comprehensive capabilities in optimizing business processes, restructuring business models, and innovating value creation mechanisms [16, 17]. In the process through which new quality productive forces enhance firm performance, digital capability plays a crucial mediating role. The cultivation and improvement of digital capability are not only intrinsic requirements for the development of new quality productive forces but also important pathways for realizing their value transformation.

Driven by technological innovation as its core engine, the development of new quality productive forces systematically promotes the enhancement of firms' digital capabilities. On the one hand, the revolutionary technological breakthroughs underlying new quality productive forces require firms to continuously deepen the research, development, and application of digital technologies, thereby accelerating the construction of digital infrastructure and the formation of digital integration capabilities [18]. On the other hand, new quality productive forces emphasize the innovative allocation of production factors, encouraging firms to incorporate data as a key production factor throughout the entire production and operation process. This accelerates the deep integration of digital technologies with traditional business activities, thereby improving firms' information perception, operational efficiency, and coordination capabilities in a digital environment [19]. Therefore, improvements in the level of new quality productive forces directly drive the cultivation and upgrading of firms' digital capabilities.

The enhancement of digital capability further promotes improvements in firm performance. In the short term, digital capability helps reduce operational costs and improve efficiency by optimizing business processes and enhancing resource allocation efficiency, thereby improving profitability and operational performance. Digital product innovation, digital process innovation, and digital service innovation—key manifestations of digital capability—significantly contribute to firms' current operating performance and can be translated into higher sales revenue and returns on assets [20, 21]. In the long term, digital capability fosters business model innovation and expands new growth opportunities, providing sustained momentum for firm value creation. Meanwhile, digital technologies enhance firms' ability to cope with market uncertainty, improving operational resilience and long-term sustainability, which ultimately manifests in continuous improvements in firm value [22].

Based on the above analysis, this study proposes the following hypothesis:

H2: New quality productive forces enhance firm performance by improving firms' digital capability.

2.3 New Quality Productive Forces, R&D Investment, and Firm Performance

R&D investment serves as the material foundation for firms' technological innovation activities and constitutes a core resource supporting technological breakthroughs and industrial upgrading. In the process of enhancing firm performance, R&D investment provides a crucial transmission channel. The technological breakthroughs and innovative allocation of production factors embodied in new quality productive forces can only be transformed into actual productive capacity through sustained R&D investment.

With technological innovation as its core, the development of new quality productive forces inevitably promotes continuous increases in firms' R&D investment. From the perspective of capability support, the efficiency improvements and value creation brought about by new quality productive forces provide stronger profit guarantees for firms to expand R&D investment, enabling them to allocate sufficient resources to technological advancement [23]. From the perspective of development momentum, the optimization of factor combinations and deep industrial upgrading driven by new quality productive forces generate sustained endogenous motivation for firms to transform technological breakthroughs into sustainable competitive advantages through R&D investment [23]. Therefore, improvements in new quality productive forces jointly promote increased R&D investment from both resource availability and strategic orientation, making technological innovation an endogenous driver of firm development.

R&D investment is a key driving factor that enables firms to convert innovation resources into performance outcomes. Continuous R&D investment can significantly enhance firms' innovation performance, particularly in industries such as manufacturing that require long-term technological accumulation, where the effect of

R&D investment on firm performance is more pronounced [24]. Cross-country studies also indicate that R&D-intensive firms enjoy clear advantages in profitability, and high-quality patent portfolios can be effectively transformed into economic returns for firms [25]. More fundamentally, sustained R&D investment enables firms to evolve from mere market participants into technology leaders. By establishing inimitable technological barriers and knowledge assets, firms generate forward-looking endogenous growth momentum, thereby achieving sustained performance growth in a highly competitive market environment [26].

Based on the above analysis, this study proposes the following hypothesis:

H3: New quality productive forces promote improvements in firm performance by increasing firms' R&D investment.

3. Research Design

3.1 Sample Selection and Data Sources

This study uses Chinese A-share listed firms in the Shanghai and Shenzhen stock markets from 2015 to 2024 as the initial sample. The sample is processed according to the following criteria: (1) observations with missing data or significant errors are excluded; (2) firms labeled as ST or *ST, as well as firms in the financial industry, are excluded. After these procedures, a total of 24,253 observations are obtained. To mitigate the influence of extreme values, all continuous variables are winsorized at the 1% and 99% levels. Data for core variables—including new quality productive forces, firm performance, and control variables—are obtained from the CSMAR and Wind databases.

3.2 Variable Definitions

3.2.1 Independent Variable

To measure firms' new quality productive forces, this study draws on the approaches of Song Jia et al. [27] and Zhang Xiu'e et al. [28]. Based on data availability, the evaluation indicators are integrated and adjusted to construct a firm-level evaluation index system for new quality productive forces, as shown in Table 1. The entropy method is then employed to calculate the variable "new quality productive forces."

Table 1: Evaluation Index System of New Quality Productive Forces at the Firm Level

Primary Dimension	Secondary Dimension	Tertiary Indicator	Measurement
New-type labor	Employee quality	Proportion of R&D personnel	$(\text{Number of R\&D personnel} / \text{Total employees}) \times 100$
		Proportion of highly educated personnel	$(\text{Number of employees with postgraduate degrees or above} / \text{Total employees}) \times 100$
	Managerial quality	Green awareness of executives	$\ln(\text{frequency of green development keywords in annual reports} + 1)$
		Overseas background of executives	Equals 1 if any executive has overseas experience; otherwise 0
New-type labor objects	Ecological environment	Environmental governance score	E score of Huazheng ESG rating, assigned values from 1 to 9
	Future development	Fixed asset ratio	$(\text{Fixed assets} / \text{Total assets}) \times 100$
		Capital accumulation rate	$(\text{Increase in owners' equity during the year} / \text{Owners' equity at the beginning of the year}) \times 100$
New-type means of production	Technological means	Innovation level	$\ln(\text{number of granted patents} + 1)$
	Digital means	Degree of digitalization	$\ln(\text{frequency of digital-related keywords in annual reports} + 1)$
		Intangible asset ratio	$(\text{Intangible assets} / \text{Total assets}) \times 100$
	Green means	Green technology level	$\ln(\text{number of granted green patents} + 1)$
		Proportion of green patents	$(\text{Number of granted green patents} / \text{Total granted patents}) \times 100$

3.2.2 Dependent Variables

Firm performance should not only be reflected in short-term operating capability but also encompass long-term value creation capacity, including future development, technological innovation, and managerial improvement. Analyzing firm performance from both long-term value and short-term operations not only aligns with the multidimensional evaluation requirements in theory but is also consistent with dynamic capability theory, which emphasizes firms' development paths in opportunity recognition, resource integration, and organizational reconfiguration. This perspective helps to more clearly reveal the evolutionary logic from cognitive transformation and capability accumulation to performance outcomes.

Long-term value performance. This study uses total factor productivity (TFP_LP) as the proxy. Total factor productivity not only captures the driving effects of technological breakthroughs and innovative allocation of production factors but also reflects the systemic improvements arising from the optimized combination of advanced labor and new-type means of production. It is therefore highly consistent with firms' long-term strategic goals of achieving high-quality development. In the context of national policies promoting the development of new quality productive forces, TFP, due to its sensitivity to technological innovation, factor reallocation, and industrial upgrading, can accurately reflect the substantive performance improvements achieved by firms under policy support. Hence, it serves as an important indicator for measuring the long-term value performance of manufacturing firms [29, 30]. The LP method and OP method are two commonly used approaches for estimating TFP. This study adopts the LP method to calculate TFP and uses it as the core indicator in the baseline regression.

Short-term operating performance. This study uses return on equity (ROE) as the proxy. ROE directly reflects a firm's ability to generate net profits from its own capital and serves as an important reference for investors in evaluating short-term profitability as well as for managerial decision-making. In the process of transformation and upgrading driven by new quality productive forces, whether through optimization of production processes or restructuring of organizational systems, the outcomes are ultimately reflected in ROE. Previous literature commonly employs indicators such as ROE and return on assets (ROA) to measure firm operating performance, as they comprehensively capture firms' overall profitability [31].

3.2.3 Mechanism Variables

Based on the preceding theoretical analysis, this study selects two mechanism variables. First, digital capability, following Huang Bo et al. [32], is measured as the natural logarithm of the number of digital patent applications plus one. The cultivation of new quality productive forces relies on technological innovation capability, and core technology patents directly reflect firms' innovation capacity. These patents provide technological support for the development of new quality productive forces and reduce the risk of mismatch between innovation input and output. Core technology patents not only represent technological advancement but also reflect firms' proactive attitude and practical engagement in developing new quality productive forces, as well as their ability to optimize resource allocation and mobilize high-quality human capital. Through the continuous accumulation and application of core technology patents, firms can gradually establish a comprehensive system of new quality productive forces and gain sustained competitive advantages in a highly competitive market. Second, R&D investment, following Guo Tianyong and Sun Guangyu [33], is measured as the ratio of R&D expenditure to operating revenue. This indicator more accurately reflects firms' emphasis on technological innovation and the intensity of resource input. R&D investment constitutes the material foundation for firms' innovation activities and provides continuous resource support for the development of new quality productive forces. Such investment not only involves financial inputs but also drives the aggregation of high-quality R&D personnel, the intelligent upgrading of production equipment, and the continuous optimization of production processes. Through sustained and stable R&D investment, firms can effectively allocate innovation resources to key core technological areas, thereby reducing the risk of disconnection between technological achievements and market applications. Compared with output-based indicators such as patent applications, R&D investment, as a leading input variable, can better signal firms' future potential for technological breakthroughs.

3.2.4 Control Variables

Drawing on the approaches of Li Dingzhao and Bao Yarong, Zhang Zhiwei et al., and Tao Xiaolong et al. [34-36], this study includes variables that may affect firms' new quality productive forces as control variables.

Specifically, eight firm-level control variables are selected: (1) **Firm basic characteristics**: listing age (ListAge) and prior loss (Loss); (2) **Financial characteristics**: firm size (Size), leverage (Lev), liquidity ratio (Liquid), and fixed asset ratio (Fixed); (3) **Corporate governance characteristics**: board size (Board) and proportion of independent directors (Indep).

The definitions and measurements of all variables are presented in Table 2.

Table 2: Variable Definitions

	Variable	VARIABLES	Measurement
Dependent variables	Long-term value performance	TFP_LP	Estimated total factor productivity using the LP method
	Short-term operating performance	ROE	Net profit / Total equity
Independent variable	New quality productive forces	Npro	Calculated using the entropy method
Mechanism variables	Digital capability	Digital	ln (number of digital patent applications + 1)
	R&D investment	R&D	R&D expenditure / Operating revenue
Control variables	Firm size	Size	Natural logarithm of total assets
	Leverage	Lev	Year-end total liabilities / Year-end total assets
	Liquidity ratio	Liquid	Current assets / Current liabilities
	Listing age	ListAge	ln (current year – year of establishment + 1)
	Prior loss	Loss	Equals 1 if net profit in the previous year < 0; otherwise 0
	Board size	Board	Natural logarithm of number of directors
	Proportion of independent directors	Indep	Number of independent directors / Total number of directors
	Fixed asset ratio	Fixed	Net fixed assets / Total assets

3.3 Model Specification

To test the proposed hypotheses, this study constructs the following models:

$$TFP\_LP_{i,t} = \alpha_0 + \alpha_1 Npro_{i,t} + \Sigma \gamma Controls_{i,t} + \delta_{stkcd} + \mu_{year} + \varepsilon_{i,t} \quad (1)$$

$$ROE_{i,t} = \alpha_0 + \alpha_1 Npro_{i,t} + \Sigma \gamma Controls_{i,t} + \delta_{stkcd} + \mu_{year} + \varepsilon_{i,t} \quad (2)$$

$$R\&D_{i,t} = \alpha_0 + \alpha_1 Npro_{i,t} + \Sigma \gamma Controls_{i,t} + \delta_{stkcd} + \mu_{year} + \varepsilon_{i,t} \quad (3)$$

$$Dig_{i,t} = \alpha_0 + \alpha_1 Npro_{i,t} + \Sigma \gamma Controls_{i,t} + \delta_{stkcd} + \mu_{year} + \varepsilon_{i,t} \quad (4)$$

In the models, subscript i denotes the firm (identified by stock code), and t denotes the year. $TFP\_LP_{i,t}$ represents the total factor productivity of firm i in year t , while $ROE_{i,t}$ denotes the return on equity of firm i in year t . $R\&D_{i,t}$ refers to the R&D investment of firm i in year t , and $Dig_{i,t}$ captures the digital capability of firm i in year t . $Controls_{i,t}$ represents a set of control variables for firm i in year t . δ denotes firm fixed effects, μ represents year fixed effects, and ε is the random error term. According to the hypotheses proposed above, the coefficients α_1 in Models (1) and (2) are expected to be positive. Models (3) and (4) are used to test the underlying mechanisms.

4. Empirical Analysis

4.1 Descriptive Statistics

Table 3 reports the descriptive statistics. The core explanatory variable, new quality productive forces (Npro), has a mean of 12.9676, a standard deviation of 6.5732, a minimum of 2.5158, and a maximum of 33.1242, indicating substantial variation in the level of new quality productive forces across firms. Regarding the dependent variables, return on equity (ROE) has a mean of 0.0543, with a minimum of -0.5294 and a maximum of 0.2990, suggesting considerable heterogeneity in profitability among sample firms, with some

firms experiencing substantial losses. Total factor productivity (TFP_LP) has a mean of 9.0615 and a standard deviation of 1.0517, with values ranging from 6.9637 to 11.9975. The relatively concentrated distribution indicates limited variation in production efficiency across firms.

At the level of control variables, firm size (Size) ranges from 20.1934 to 26.3264, with a mean of 22.3813, indicating a relatively even distribution of firm sizes. Leverage (Lev) ranges from 0.0616 to 0.8590, suggesting significant variation in capital structure across firms, with some firms having low financial leverage while others face higher debt pressure. The liquidity ratio (Liquid) ranges from 0.3487 to 13.9741, indicating that some firms have relatively weak short-term solvency while others maintain ample liquidity. Listing age (ListAge) ranges from 0.6931 to 3.4340, with a mean of 2.2245, showing a relatively concentrated distribution. The dummy variable for prior loss (Loss) has a mean of 0.1021, indicating that approximately 10.21% of sample firms reported losses in the previous year. Board size (Board) ranges from 1.6094 to 2.6391, with a mean of 2.1043, reflecting variation in board size across firms. The proportion of independent directors (Indep) ranges from 33.33% to 57.14%, with a mean of 37.80%, generally complying with regulatory requirements. The fixed asset ratio (Fixed) ranges from 0.0029 to 0.6499, with a mean of 0.2093, indicating significant differences in asset structure across firms.

Table 3: Descriptive Statistics of Variables

VARIABLES	N	Min	Max	Mean	p50	SD
Npro	24253	2.5158	33.1242	12.9676	12.4775	6.5732
ROE	24253	-0.5294	0.2990	0.0543	0.0636	0.1108
TFP_LP	24253	6.9637	11.9975	9.0615	8.9362	1.0517
Size	24253	20.1934	26.3264	22.3813	22.1914	1.2597
Lev	24253	0.0616	0.8590	0.4129	0.4068	0.1930
Liquid	24253	0.3487	13.9741	2.4313	1.7285	2.1854
ListAge	24253	0.6931	3.4340	2.2245	2.3026	0.7835
Loss	24253	0.0000	1.0000	0.1021	0.0000	0.3028
Board	24253	1.6094	2.6391	2.1043	2.1972	0.1944
Indep	24253	0.3333	0.5714	0.3780	0.3636	0.0529
Fixed	24253	0.0029	0.6499	0.2093	0.1820	0.1471

4.2 Baseline Regression

The baseline regression results are reported in Table 4. Column (1) uses return on equity (ROE) as a proxy for firms' short-term operating performance, while column (2) uses total factor productivity (TFP_LP) as a proxy for long-term value performance. Both models include control variables and are estimated using a fixed-effects model. The results in column (1) show that the coefficient of new quality productive forces (Npro) is 0.0005 and is significantly positive at the 1% level, indicating that new quality productive forces significantly enhance firms' short-term profitability. Column (2) shows that the coefficient of Npro on TFP_LP is 0.0016, also significantly positive at the 1% level, suggesting that new quality productive forces positively affect firms' production efficiency. These findings indicate that new quality productive forces improve firm performance by promoting technological innovation, optimizing resource allocation, and enhancing production efficiency.

The adjusted R^2 of both models is 0.936, indicating strong explanatory power. Overall, the baseline regression results demonstrate that new quality productive forces have a significantly positive impact on both short-term operating performance and long-term value performance, thereby providing initial support for Hypothesis H1.

Table 4: Baseline Regression Results

VARIABLES	(1)	(2)
	ROE	TFP_LP
Npro	0.0005*** (3.3840)	0.0016*** (3.2864)
Size	0.0513*** (24.2075)	0.6054*** (92.5377)
Lev	-0.3197*** (-36.5194)	-0.0268 (-0.9925)
Liquid	-0.0082***	-0.0267***

VARIABLES	(1)	(2)
	ROE	TFP LP
	(-12.2938)	(-12.8879)
ListAge	-0.0237***	-0.0212**
	(-7.0934)	(-2.0568)
Loss	-0.0258***	-0.0740***
	(-11.5765)	(-10.7467)
Board	-0.0053	-0.0222
	(-0.6456)	(-0.8807)
Indep	-0.0269	-0.0092
	(-1.0977)	(-0.1220)
Fixed	-0.1728***	-0.9923***
	(-16.1852)	(-30.1044)
Constant	-0.8362***	-4.1215***
	(-16.6795)	(-26.6352)
Firm Fixed	YES	YES
Year Fixed	YES	YES
Observations	24,253	24,253
R-squared	0.488	0.946
Adj. R ²	0.936	0.936

*Note: *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively; t-values are reported in parentheses. The same applies below.

4.3 Mechanism Analysis

To examine the mechanisms through which R&D investment and digital capability affect the relationship between new quality productive forces and firm performance, this study follows Jiang Ting [37] and conducts mechanism tests. The results are presented in Table 5. Column (1) uses R&D intensity (R&D) as the mechanism variable, while column (2) uses digital capability (Digital). Both regressions include control variables.

The results in column (1) show that the coefficient of Npro is 0.0001 and is significantly positive at the 1% level, indicating that improvements in new quality productive forces significantly promote firms' R&D investment intensity. Column (2) shows that the coefficient of Npro on digital capability is 0.0172, also significantly positive at the 1% level, suggesting that new quality productive forces significantly enhance firms' digital capability. These findings indicate that new quality productive forces, on the one hand, strengthen firms' R&D investment intensity by promoting technological innovation and R&D activities; on the other hand, they accelerate digital transformation and improve digital capability. Therefore, Hypotheses H2 and H3 are supported.

Table 5: Mechanism Test Results

VARIABLES	(1)	(2)
	R&D	Digital
Npro	0.0001***	0.0172***
	(3.7420)	(13.8567)
Size	-0.0024***	0.4826***
	(-4.7644)	(27.8581)
Lev	-0.0142***	-0.0868
	(-6.7683)	(-1.2146)
Liquid	0.0002	-0.0166***
	(1.5019)	(-3.0816)
ListAge	0.0034***	-0.1280***
	(4.2472)	(-4.7423)
Loss	0.0015***	-0.0335*
	(2.8771)	(-1.8653)
Board	0.0028	-0.0217
	(1.4435)	(-0.3318)
Indep	0.0029	0.0295

VARIABLES	(1)	(2)
	R&D	Digital
	(0.5084)	(0.1507)
Fixed	0.0185***	0.0885
	(7.0768)	(0.9920)
Constant	0.0907***	-8.3506***
	(7.5634)	(-20.4233)
Firm Fixed	YES	YES
Year Fixed	YES	YES
Observations	22,256	22,256
R-squared	0.874	0.878
Adj. R ²	0.856	0.856

4.4 Endogeneity Tests

4.4.1 Instrumental Variable Approach

To mitigate potential endogeneity issues in the baseline regression—such as reverse causality and omitted variable bias—this study employs an instrumental variable approach and conducts two-stage least squares (2SLS) estimation using two instruments.

First, the development of new quality productive forces is typically gradual and path-dependent, and its impact on firm performance may involve a time lag. Following Li Hongkai [38], this study uses the one-period lag of new quality productive forces (L.Npro) as an instrumental variable. Given the persistence and path dependence of new quality productive forces, the lagged term is highly correlated with the current level of Npro, while its direct impact on current firm performance is limited after controlling for firm and year fixed effects. Therefore, it helps alleviate potential endogeneity concerns. The first-stage regression results are reported in Table 6. In the model using a one-period lagged instrumental variable, the coefficient of the instrumental variable is 0.7319 and is significantly positive at the 1% level, indicating that the instrument is valid and does not suffer from weak instrument problems. In the second-stage regression, the fitted values of Npro obtained from the instrumental variable remain significantly positive at the 1% level for both ROE and TFP_LP, indicating that the conclusion that “new quality productive forces significantly promote firm performance” remains robust after addressing endogeneity issues.

Table 6: Instrumental Variable Results (Lagged Variable)

VARIABLES	(1)	(2)	(3)	(4)
	first	second	first	second
	Npro	ROE	Npro	TFP_LP
LNpro	0.7319***		0.7319***	
	(154.4275)		(154.4275)	
Npro		0.0012***		0.0125***
		(7.5641)		(15.1581)
Size	0.4072***	0.0267***	0.4072***	0.6547***
	(13.0176)	(34.6988)	(13.0176)	(161.5582)
Lev	-0.4062*	-0.1985***	-0.4062*	0.3748***
	(-1.7253)	(-34.9578)	(-1.7253)	(12.5245)
Liquid	-0.0401**	-0.0044***	-0.0401**	-0.0362***
	(-1.9894)	(-8.9580)	(-1.9894)	(-14.1043)
ListAge	-0.6176***	-0.0170***	-0.6176***	0.0151**
	(-12.5933)	(-14.1837)	(-12.5933)	(2.3836)
Loss	-0.2314**	-0.0949***	-0.2314**	-0.1848***
	(-2.2978)	(-39.0808)	(-2.2978)	(-14.4396)
Board	0.2123	0.0007	0.2123	-0.1489***
	(1.0755)	(0.1519)	(1.0755)	(-5.9373)
Indep	0.3708	0.0160	0.3708	-0.2352***
	(0.5399)	(0.9689)	(0.5399)	(-2.6955)
Fixed	-1.0608***	-0.0301***	-1.0608***	-1.1186***
	(-4.9867)	(-5.8399)	(-4.9867)	(-41.2270)
Constant	-7.7716***	-0.4346***	-7.7716***	-5.2071***

VARIABLES	(1)	(2)	(3)	(4)
	first	second	first	second
	Npro	ROE	Npro	TFP_LP
	(-10.0473)	(-23.1541)	(-10.0473)	(-52.6323)
Firm Fixed	YES	YES	YES	YES
Year Fixed	YES	YES	YES	YES
Observations	19,309	19,309	19,309	19,309
R-squared	0.602	0.224	0.602	0.754
Adj. R ²	0.601	0.224	0.601	0.754

Second, following Su Taoyong et al. [39], this study constructs another instrumental variable: the mean level of new quality productive forces for firms within the same year–province–industry group (meanNpro). Given the presence of spillover effects, a firm's level of new quality productive forces is correlated with that of other firms in the same group, while the latter is unlikely to directly affect the focal firm's performance. Thus, this variable satisfies both the relevance and exogeneity conditions of a valid instrument. In the mean-based IV model reported in Table 7, the coefficient of the instrumental variable in the first stage is 0.9712 and is significantly positive at the 1% level, indicating no under-identification or weak instrument problems. In the second stage, the coefficients of the instrumented Npro on ROE and TFP_LP are 0.0009 and 0.0086, respectively, both significant at the 1% level. These results further confirm that the positive effect of new quality productive forces on firm performance remains robust after controlling for endogeneity.

Table 7: Instrumental Variable Results (Mean-Based Instrument)

VARIABLES	(1)	(2)	(3)	(4)
	first	second	first	second
	Npro	ROE	Npro	TFP_LP
meanNpro	0.9712*** (124.4421)		0.9712*** (124.4421)	
Npro		0.0009*** (5.3154)		0.0086*** (9.9214)
Size	0.9263*** (28.2248)	0.0260*** (37.7563)	0.9263*** (28.2248)	0.6566*** (177.2326)
Lev	-1.0029*** (-4.0285)	-0.1905*** (-38.2400)	-1.0029*** (-4.0285)	0.3661*** (13.6418)
Liquid	-0.0745*** (-3.6825)	-0.0045*** (-11.0666)	-0.0745*** (-3.6825)	-0.0379*** (-17.4067)
ListAge	-0.5602*** (-12.0251)	-0.0165*** (-17.2767)	-0.5602*** (-12.0251)	0.0052 (1.0049)
Loss	-0.0046 (-0.0419)	-0.0938*** (-43.2581)	-0.0046 (-0.0419)	-0.1947*** (-16.6601)
Board	0.8660*** (4.1195)	-0.0005 (-0.1255)	0.8660*** (4.1195)	-0.1437*** (-6.3476)
Indep	0.2881 (0.3912)	0.0099 (0.6727)	0.2881 (0.3912)	-0.2363*** (-2.9815)
Fixed	-1.4191*** (-6.3176)	-0.0395*** (-8.7543)	-1.4191*** (-6.3176)	-1.1490*** (-47.2169)
Constant	-20.0799*** (-24.3259)	-0.3960*** (-23.7667)	-20.0799*** (-24.3259)	-5.2313*** (-58.2801)
Firm Fixed	YES	YES	YES	YES
Year Fixed	YES	YES	YES	YES
Observations	24,542	24,542	24,542	24,542
R-squared	0.447	0.218	0.447	0.749
Adj. R ²	0.447	0.217	0.447	0.749

4.4.2 Propensity Score Matching (PSM)

Firms with different levels of new quality productive forces may also differ significantly in other aspects, which could affect firm performance and lead to sample self-selection bias. Following Guo Wei and Kang Wenyi [40], this study employs propensity score matching (PSM) to mitigate this endogeneity issue. Specifically, (1) the sample is divided into a treatment group and a control group based on the median value

of new quality productive forces; (2) the eight control variables used in this study are treated as covariates, and kernel matching is applied to match the treatment and control groups, ensuring that the matched samples exhibit no significant differences in observable characteristics other than new quality productive forces; (3) the main effects are re-estimated using the matched sample, with the results reported in Table 8. The regression coefficient of new quality productive forces (Npro) on ROE is 0.0006, and that on TFP_LP is 0.0013, both remaining significantly positive. These results further confirm the robustness of the baseline regression findings.

Table 8: Propensity Score Matching Results

VARIABLES	(1)	(2)
	ROE	TFP_LP
Npro	0.0006*** (3.2497)	0.0013** (2.4662)
Size	0.0543*** (21.8577)	0.5952*** (77.1901)
Lev	-0.3099*** (-31.2856)	-0.0311 (-1.0134)
Liquid	-0.0088*** (-11.5134)	-0.0284*** (-12.0246)
ListAge	-0.0275*** (-7.4591)	-0.0182 (-1.5943)
Loss	-0.0251*** (-10.0296)	-0.0708*** (-9.1223)
Board	0.0014 (0.1508)	-0.0269 (-0.9554)
Indep	-0.0388 (-1.4272)	0.0606 (0.7197)
Fixed	-0.1655*** (-13.9879)	-0.9771*** (-26.6196)
Constant	-0.9078*** (-15.5006)	-3.9213*** (-21.5850)
Firm Fixed	YES	YES
Year Fixed	YES	YES
Observations	19,681	19,681
R-squared	0.481	0.941
Adj. R ²	0.929	0.929

4.5 Robustness Tests

First, excluding specific years. Considering that the COVID-19 pandemic may have had a potential impact on the relationship between new quality productive forces and firm performance, this study excludes observations from 2019 to 2021 to eliminate pandemic-related disturbances and re-estimates the model. As shown in Panel A of Table 9, the coefficient of Npro remains significantly positive at the 1% level. This indicates that the positive effect of new quality productive forces on firm performance remains robust even after excluding the potential impact of the COVID-19 shock. These results further confirm the robustness of the findings.

Second, clustered robust standard errors. After incorporating firm and year fixed effects in the baseline regression, standard errors are clustered at the firm level to account for within-firm correlation. The results, reported in Panel B of Table 9, show that the coefficient of the core explanatory variable remains significantly positive, indicating that the baseline results are robust.

Third, adding province fixed effects. To control for potential regional differences in new quality productive forces, province fixed effects are further included in the regression. As shown in Panel C of Table 9, the coefficients of Npro are 0.005 and 0.0017, both significantly positive, consistent with the baseline results and confirming the robustness of the conclusions.

Table 9: Robustness Test Results

VARIABLES	Panel A: Excluding Specific Years		Panel B: Clustered Robust Standard Errors		Panel C: Adding Province Fixed Effects	
	(1)	(2)	(1)	(2)	(1)	(2)
	ROE	TFP LP	ROE	TFP LP	ROE	TFP LP
Npro	0.0006*** (3.0966)	0.0016*** (2.5878)	0.0005*** (3.5818)	0.0016*** (3.2519)	0.0005*** (3.3317)	0.0017*** (3.5091)
Size	0.0444*** (18.7567)	0.6028*** (78.0395)	0.0513*** (15.4796)	0.6054*** (55.4247)	0.0497*** (23.4136)	0.6052*** (92.4296)
Lev	-0.2923*** (-28.4391)	-0.0085 (-0.2523)	-0.3197*** (-23.0483)	-0.0268 (-0.7028)	-0.3182*** (-36.3444)	-0.0253 (-0.9378)
Liquid	-0.0080*** (-9.9346)	-0.0246*** (-9.4157)	-0.0082*** (-12.5761)	-0.0267*** (-9.7114)	-0.0084*** (-12.5649)	-0.0264*** (-12.7588)
ListAge	-0.0240*** (-6.3884)	-0.0127 (-1.0365)	-0.0237*** (-7.5798)	-0.0212* (-1.8483)	-0.0238*** (-7.1219)	-0.0218** (-2.1194)
Loss	-0.0323*** (-11.9352)	-0.0847*** (-9.5650)	-0.0258*** (-7.6427)	-0.0740*** (-8.7282)	-0.0252*** (-11.3093)	-0.0739*** (-10.7617)
Board	-0.0141 (-1.4478)	-0.0805** (-2.5375)	-0.0053 (-0.5294)	-0.0222 (-0.7910)	-0.0052 (-0.6319)	-0.0265 (-1.0533)
Indep	-0.0405 (-1.3883)	-0.1329 (-1.3939)	-0.0269 (-0.9548)	-0.0092 (-0.1112)	-0.0237 (-0.9704)	-0.0344 (-0.4572)
Fixed	-0.1758*** (-13.9414)	-0.9529*** (-23.1322)	-0.1728*** (-13.5559)	-0.9923*** (-20.7232)	-0.1749*** (-16.3940)	-0.9821*** (-29.8665)
Constant	-0.6705*** (-11.9035)	-3.9480*** (-21.4579)	-0.8362*** (-11.3184)	-4.1215*** (-16.6809)	-0.8019*** (-15.9501)	-4.1006*** (-26.4612)
Firm Fixed	YES	YES	YES	YES	YES	YES
Year Fixed	YES	YES	YES	YES	YES	YES
Province Fixed	NO	NO	NO	NO	YES	YES
Observations	16,667	16,667	24,253	24,253	24,253	24,253
R-squared	0.523	0.945	0.488	0.946	0.491	0.946
Adj. R ²	0.930	0.930	0.936	0.936	0.937	0.937

4.6 Heterogeneity Analysis

Significant differences exist in the levels of new quality productive forces and firm performance across firms, regions, and industries. Accordingly, the economic effects of developing new quality productive forces may exhibit substantial heterogeneity. This study conducts subgroup analyses to examine heterogeneity in the impact of new quality productive forces on firm performance, based on whether firms belong to heavily polluting industries, regional economic development levels, and geographic location.

4.6.1 Heterogeneity across Heavily Polluting Industries

The sample is divided into two groups: heavily polluting industries and non-heavily polluting industries. Separate regressions are conducted for each group, with results reported in Table 10. For firms in non-heavily polluting industries, the coefficients of new quality productive forces on short-term operating performance and long-term value performance are 0.0006 and 0.0015, respectively, both significant at the 1% level. In contrast, for firms in heavily polluting industries, the effects of new quality productive forces on both performance measures are not statistically significant. These results indicate that firms in non-heavily polluting industries can significantly enhance performance through the development of new quality productive forces, whereas firms in heavily polluting industries do not experience significant improvements in short-term performance, nor significant changes in long-term performance. Firms in heavily polluting industries generally face higher environmental regulatory pressure and pollution control costs, rely more heavily on traditional production factors, and exhibit stronger path dependence in technological innovation and digital transformation. As a result, they must simultaneously address environmental compliance and technological upgrading when developing new quality productive forces, facing higher transformation costs and institutional constraints. By contrast, firms in non-heavily polluting industries possess greater flexibility and innovation space in resource

allocation and technology application, enabling them to more effectively translate new quality productive forces into improved operating performance.

Table 10: Heterogeneity Analysis: Heavily Polluting vs. Non-Heavily Polluting Industries

VARIABLES	Heavily Polluting Industries		Non-Heavily Polluting Industries	
	(1)	(2)	(3)	(4)
	ROE	TFP_LP	ROE	TFP_LP
Npro	0.0003	-0.0002	0.0006***	0.0015***
	(0.7207)	(-0.1267)	(3.3525)	(2.8456)
Size	0.0526***	0.5334***	0.0504***	0.6079***
	(6.9739)	(27.3957)	(22.0662)	(85.6152)
Lev	-0.3543***	-0.4525***	-0.3118***	0.0294
	(-13.3393)	(-6.5983)	(-33.5269)	(1.0161)
Liquid	-0.0095***	-0.0380***	-0.0080***	-0.0257***
	(-4.1669)	(-6.4224)	(-11.4276)	(-11.7724)
ListAge	-0.0332***	0.0113	-0.0203***	-0.0199*
	(-2.8405)	(0.3753)	(-5.7650)	(-1.8107)
Loss	-0.0293***	-0.0579***	-0.0248***	-0.0752***
	(-4.5275)	(-3.4686)	(-10.4257)	(-10.1568)
Board	0.0028	0.0138	-0.0049	-0.0302
	(0.1117)	(0.2106)	(-0.5644)	(-1.1235)
Indep	0.0530	0.1468	-0.0358	-0.0331
	(0.7597)	(0.8151)	(-1.3682)	(-0.4068)
Fixed	-0.1456***	-0.8570***	-0.1807***	-1.0238***
	(-5.0636)	(-11.5426)	(-15.4711)	(-28.1998)
Constant	-0.8676***	-2.3770***	-0.8249***	-4.1865***
	(-4.7590)	(-5.0493)	(-15.3659)	(-25.0954)
Firm Fixed	YES	YES	YES	YES
Year Fixed	YES	YES	YES	YES
Observations	2,492	2,492	21,742	21,742
R-squared	0.481	0.962	0.491	0.945
Adj. R ²	0.935	0.935	0.935	0.935

4.6.2 Heterogeneity by Economic Development Level

This study uses the logarithm of regional GDP per capita as a proxy for the level of economic development and divides the sample into two groups: regions with high economic development and regions with low economic development. The regression results are reported in Table 11. For the high economic development group, the coefficients of new quality productive forces on short-term operating performance (ROE) and long-term value performance (TFP_LP) are 0.0007 and 0.0021, respectively, both significant at the 1% level. In contrast, for the low economic development group, the effects of new quality productive forces on both performance measures are not statistically significant. These results indicate that in regions with higher levels of economic development, firms can effectively improve their performance by developing new quality productive forces, whereas in less developed regions, this promoting effect has not yet materialized. Regions with higher levels of economic development typically possess more advanced digital infrastructure, a stronger concentration of innovation factors, and more mature market mechanisms, enabling firms to more effectively translate new quality productive forces into actual outputs. In contrast, regions with lower levels of economic development are relatively constrained in terms of technological absorption capacity, human capital endowment, and industrial supporting systems. As a result, firms in these regions face higher technological barriers and institutional frictions in promoting the transformation toward new quality productive forces, which weakens their positive impact on firm performance.

Table 11: Heterogeneity Analysis by Economic Development Level

VARIABLES	High Economic Development		Low Economic Development	
	(1)	(2)	(3)	(4)
	ROE	TFP_LP	ROE	TFP_LP
Npro	0.0007***	0.0021***	0.0002	0.0008
	(3.3100)	(3.0989)	(0.7102)	(1.1601)
Size	0.0514***	0.5843***	0.0540***	0.6311***

VARIABLES	High Economic Development		Low Economic Development	
	(1)	(2)	(3)	(4)
	ROE	TFP LP	ROE	TFP LP
	(17.4206)	(63.2642)	(17.0878)	(66.0629)
Lev	-0.2875***	0.0589	-0.3611***	-0.1765***
	(-23.6864)	(1.5499)	(-27.8051)	(-4.4964)
Liquid	-0.0086***	-0.0238***	-0.0079***	-0.0294***
	(-9.5833)	(-8.5176)	(-7.6359)	(-9.4017)
ListAge	-0.0216***	-0.0107	-0.0265***	-0.0109
	(-4.7976)	(-0.7595)	(-5.0234)	(-0.6819)
Loss	-0.0176***	-0.0668***	-0.0260***	-0.0698***
	(-5.6197)	(-6.7963)	(-8.0206)	(-7.1297)
Board	0.0015	-0.0204	-0.0067	-0.0023
	(0.1322)	(-0.5633)	(-0.5661)	(-0.0649)
Indep	-0.0031	0.0962	-0.0373	-0.1580
	(-0.0894)	(0.8796)	(-1.0572)	(-1.4814)
Fixed	-0.1851***	-1.1239***	-0.1427***	-0.8027***
	(-12.2221)	(-23.7165)	(-9.2712)	(-17.2541)
Constant	-0.8831***	-3.7259***	-0.8645***	-4.6856***
	(-12.5491)	(-16.9228)	(-11.5933)	(-20.7896)
Firm Fixed	YES	YES	YES	YES
Year Fixed	YES	YES	YES	YES
Observations	12,434	12,434	11,631	11,631
R-squared	0.521	0.950	0.497	0.947
Adj. R ²	0.938	0.938	0.938	0.938

4.6.3 Regional Heterogeneity

The sample is further divided into three groups: eastern, central, and western regions. Separate regressions are conducted for each group, with results reported in Table 12. The results show that the coefficient of new quality productive forces is significantly positive in the eastern region, whereas it is not statistically significant in the central and western regions. This suggests that the development of new quality productive forces significantly enhances firm performance in the eastern region, while its effect remains insignificant in the central and western regions. This pattern can be attributed to the more mature level of marketization in the eastern region, as well as its more favorable industrial structure and innovation environment, which facilitate the cultivation and transformation of new quality productive forces into actual performance gains. By contrast, the central and western regions are relatively disadvantaged in terms of infrastructure, human capital, and industrial support systems. Firms in these regions may face higher transformation costs and market frictions when promoting new quality productive forces, thereby limiting their performance-enhancing effects.

Table 12: Regional Heterogeneity Analysis

VARIABLES	Eastern Region		Central Region		Western Region	
	(1)	(2)	(3)	(4)	(5)	(6)
	ROE	TFP LP	ROE	TFP LP	ROE	TFP LP
Npro	0.0007***	0.0022***	0.0002	-0.0005	-0.0003	0.0014
	(3.8776)	(3.8646)	(0.5086)	(-0.3742)	(-0.6148)	(0.9558)
Size	0.0562***	0.6118***	0.0409***	0.5487***	0.0335***	0.6034***
	(22.5385)	(81.3427)	(6.9105)	(30.1386)	(5.9046)	(30.6841)
Lev	-0.3144***	0.0139	-0.3089***	0.1171	-0.3113***	-0.3606***
	(-30.1739)	(0.4423)	(-13.3089)	(1.6398)	(-13.7335)	(-4.5832)
Liquid	-0.0085***	-0.0227***	-0.0083***	-0.0326***	-0.0076***	-0.0456***
	(-10.8889)	(-9.5979)	(-4.7024)	(-6.0036)	(-3.9979)	(-6.8702)
ListAge	-0.0251***	-0.0271**	-0.0203**	0.0185	-0.0190**	0.0003
	(-6.4623)	(-2.3116)	(-2.0691)	(0.6139)	(-2.0589)	(0.0100)
Loss	-0.0201***	-0.0649***	-0.0391***	-0.0589***	-0.0305***	-0.1174***
	(-7.5280)	(-8.0530)	(-6.7435)	(-3.3064)	(-5.5343)	(-6.1412)
Board	0.0065	-0.0195	-0.0369*	-0.0280	-0.0034	0.0172
	(0.6749)	(-0.6754)	(-1.7015)	(-0.4200)	(-0.1514)	(0.2221)

VARIABLES	Eastern Region		Central Region		Western Region	
	(1)	(2)	(3)	(4)	(5)	(6)
	ROE	TFP LP	ROE	TFP LP	ROE	TFP LP
Indep	0.0008	0.1157	0.0009	-0.3908*	-0.1438**	-0.4181*
	(0.0272)	(1.3296)	(0.0133)	(-1.9210)	(-2.2615)	(-1.8940)
Fixed	-0.1902***	-1.0854***	-0.1685***	-0.7493***	-0.1084***	-0.8083***
	(-14.6671)	(-27.7620)	(-5.8692)	(-8.4794)	(-4.4553)	(-9.5674)
Constant	-0.9812***	-4.2998***	-0.5494***	-2.8912***	-0.4075***	-4.0125***
	(-16.6565)	(-24.2049)	(-3.9886)	(-6.8199)	(-2.9626)	(-8.4044)
Firm Fixed	YES	YES	YES	YES	YES	YES
Year Fixed	YES	YES	YES	YES	YES	YES
Observations	17,687	17,687	3,545	3,545	3,016	3,016
R-squared	0.494	0.948	0.495	0.944	0.942	0.942
Adj. R ²	0.932	0.932	0.932	0.932	0.932	0.932

5. Conclusions and Implications

Using panel data of Chinese A-share listed firms from 2015 to 2024, this study examines the impact of new quality productive forces on firms' short-term operating performance and long-term value performance, along with the underlying mechanisms, robustness, and consistency of the results. The main findings are as follows. First, new quality productive forces significantly enhance firm performance. This conclusion remains robust after excluding specific years, applying clustered robust standard errors, and altering fixed-effects specifications. Second, mechanism analysis shows that R&D investment and digital capability play important roles in the relationship between new quality productive forces and firm performance. This indicates that new quality productive forces improve firm performance by enhancing firms' innovation transformation capacity and digital technological capability. Third, heterogeneity analysis reveals that the performance-enhancing effect of new quality productive forces is more pronounced in non-heavily polluting industries and varies across regions. Specifically, the effect is stronger in regions with higher levels of economic development than in less developed regions, and it is more significant for firms in the eastern region than for those in the central and western regions.

Based on these findings, the following implications are proposed. First, firms should accelerate the cultivation of new quality productive forces, actively adapt to the new wave of technological transformation, and improve the quality of innovation. By adhering to innovation-driven development and vigorously promoting new quality productive forces, firms can effectively enhance their value creation capacity. Second, firms should recognize that R&D investment and digital capability serve as key bridges linking new quality productive forces to firm performance. On the one hand, firms should continuously increase R&D investment and establish in-depth collaborations with universities and research institutions to enhance innovation efficiency and technological breakthroughs. On the other hand, firms should accelerate digital transformation by building professional teams and improving information management systems, thereby strengthening overall digital capability and creating favorable conditions for the cultivation and application of new quality productive forces. Third, from the perspective of government policy, it is necessary to develop and improve policy systems that support the development of new quality productive forces. Particular attention should be paid to providing targeted support for firms in heavily polluting industries, strengthening assistance for environmental technology upgrading, and promoting the simultaneous achievement of green transformation and improved firm performance. In addition, given the imbalance in regional development, differentiated regional policies should be implemented to facilitate the accumulation and application of new quality productive force elements in central and western regions as well as in less developed areas, thereby promoting coordinated and high-quality regional economic development.

References

- [1] Xi, J. (2023). Xi Jinping emphasized during his inspection in Heilongjiang to firmly grasp the strategic positioning in national development and strive to create a new situation of high-quality development. *People's Daily*, 1.

- [2] Chen, J., Zhang, Y., & Wang, D. (2024). The factor mechanism and practical path of new quality productive forces empowering comprehensive rural revitalization. *Economic Review*, (4), 29–38. <https://doi.org/10.16528/j.cnki.22-1054/f.202404029>
- [3] State Council. (2024). *Government work report: Delivered at the second session of the 14th National People's Congress on March 5, 2024*. *People's Daily*, 1.
- [4] Ha, J., Lin, L., Chen, X., & Fan, Y. (2025). Network embeddedness, new quality productive forces, and firm performance in innovative industrial clusters. *Scientific Decision Making*, (9), 1–21.
- [5] Qi, Y., & Liu, J. (2025). New quality productive forces and firm value from a financial perspective: Threshold effects based on profitability levels. *Communication of Finance and Accounting*, (21), 36–41. <https://doi.org/10.16144/j.cnki.issn1002-8072.20250408.001>
- [6] Li, X., Tian, Z., & Chang, B. (2024). New quality productive forces, resource utilization, and organizational resilience. *West Forum*, 34(4), 35–49.
- [7] Zhou, W., & Xu, L. (2023). On new quality productive forces: Connotations, characteristics, and key focal points. *Reform*, (10), 1–13.
- [8] Zhang, L., & Pu, Q. (2023). Connotation, theoretical innovation, and value implications of new quality productive forces. *Journal of Chongqing University (Social Sciences Edition)*, 29(6), 137–148.
- [9] Xiao, Y., Zhang, X., & Liu, X. (2024). New quality productive forces and intra-firm pay gap: A perspective of shared development. *Economic Review*, (3), 75–91. <https://doi.org/10.19361/j.er.2024.03.05>
- [10] Bhaskar, R., Chortane, S. G., Kumar, A., & Pandey, D. K. (2026). Mandatory sustainability reporting and firm performance: A quasi-natural experiment. *Finance Research Letters*, 94, 109678. <https://doi.org/10.1016/j.frl.2026.109678>
- [11] Liu, D., Wang, X., Wang, S., & Shi, Y. (2025). New quality productive forces empowering firm resilience: An analytical framework based on new production relations and factors. *Collected Essays on Finance and Economics*, (1), 15–25. <https://doi.org/10.13762/j.cnki.cjlc.20241012.001>
- [12] Manikandan, K. S., & Balaji, V. V. (2026). Ownership form, strategy distinctiveness, and firm performance. *Asia Pacific Journal of Management*. <https://doi.org/10.1007/S10490-026-10113-0>
- [13] Cao, Y. (2024). The impact of digital new quality productive forces on industrial chain resilience. *Statistics & Decision*, 40(10), 23–27. <https://doi.org/10.13546/j.cnki.tjyjc.2024.10.004>
- [14] Sadeghi, A., Aliasghar, O., & Sadeghi, V. J. (2026). Alliances with frenemies: Capability-building mechanisms linking coopetition to firm performance. *Journal of Business Research*, 208, 116076. <https://doi.org/10.1016/j.jbusres.2026.116076>
- [15] Hu, Q. (2020). Mechanisms and performance of enterprise digital transformation. *Zhejiang Academic Journal*, (2), 146–154. <https://doi.org/10.16235/j.cnki.33-1005/c.2020.02.017>
- [16] Liu, X., Yang, Y., & Sun, Z. (2022). The construction and evolution of enterprise digital capability: A multi-case exploratory study. *Reform*, (10), 45–64.
- [17] Jiang, H., Wang, Z., Chen, C., & Gai, J. (2025). How managers' digital literacy promotes ambidextrous innovation: The serial mediating role of knowledge re-orchestration and digital capabilities. *Business Process Management Journal*, 31(8), 199–222. <https://doi.org/10.1108/BPMJ-05-2025-0672>
- [18] Sun, Z., & Lu, R. (2023). A review and prospect of enterprise digital transformation research. *Journal of Capital University of Economics and Business*, 25(6), 93–108. <https://doi.org/10.13504/j.cnki.issn1008-2700.2023.06.007>
- [19] Guan, Y., Tang, Z., Tian, M., & Du, H. (2022). The impact of digital capability on corporate entrepreneurship: The moderating role of competitive intensity. *Technology Economics*, 41(6), 95–106.
- [20] Lv, F., Zhu, Y., Robert, K., & Zhou, J. (2022). Value chain paths of digital innovation in SMEs. *Science and Technology Management Research*, 42(8), 102–110.

- [21] Hu, Y., & Sun, L. (2025). The influence of digital innovation on innovation performance in manufacturing: The mediating role of dynamic capabilities. *Technology Analysis & Strategic Management*, 37(11), 1356–1370. <https://doi.org/10.1080/09537325.2024.2309296>
- [22] Chen, S., & Wang, D. (2023). Digital transformation and firm resilience: Effects and mechanisms. *Journal of Xi'an University of Finance and Economics*, 36(4), 65–77. <https://doi.org/10.19331/j.cnki.jxufe.2023.04.005>
- [23] Wang, X., Luan, X., & Zhang, S. (2023). R&D investment, ESG performance, and market value: The moderating role of digitalization. *Studies in Science of Science*, 41(5), 896–904. <https://doi.org/10.16192/j.cnki.1003-2053.20220606.001>
- [24] Ba, S., Wu, L., & Xiong, P. (2022). Government subsidies, R&D investment, and innovation performance. *Statistics & Decision*, 38(5), 166–169. <https://doi.org/10.13546/j.cnki.tjyj.2022.05.032>
- [25] Chen, H., Qian, L., Gu, H., et al. (2025). Patent quality, R&D investment, and the profitability of technology-based firms. *Finance Research Letters*, 76, 106923.
- [26] Li, H., & Geng, X. (2021). R&D investment, executive incentives, and firm performance: Evidence from Chinese listed firms. *Journal of Harbin University of Commerce (Social Sciences Edition)*, (6), 36–48.
- [27] Song, J., Zhang, J., & Pan, Y. (2024). ESG development and new quality productive forces: Evidence from Chinese A-share firms. *Contemporary Economic Management*, 46(6), 1–11. <https://doi.org/10.13253/j.cnki.ddjjgl.2024.06.001>
- [28] Zhang, X., Wang, W., & Yu, Y. (2025). Digital-intelligent transformation and new quality productive forces. *Studies in Science of Science*, 43(5), 943–954. <https://doi.org/10.16192/j.cnki.1003-2053.20240518.003>
- [29] Shi, D., & Sun, G. (2024). Data elements and new quality productive forces: A firm-level TFP perspective. *Economic Theory and Business Management*, 44(4), 12–30.
- [30] Huang, Q., & Sheng, F. (2024). The system of new quality productive forces: Factor characteristics, structural support, and functional orientation. *Reform*, (2), 15–24.
- [31] Shi, W., Ndofor, H. A., & Hoskisson, R. E. (2020). Disciplining role of short sellers: Evidence from M&A activity. *Journal of Management*, 47(5). <https://doi.org/10.1177/0149206320912307>
- [32] Huang, B., Li, H., Liu, J., & Lei, J. (2023). Digital technological innovation and high-quality development of Chinese firms: Evidence from digital patents. *Economic Research Journal*, 58(3), 97–115.
- [33] Guo, T., & Sun, G. (2021). Economic policy uncertainty, financing costs, and corporate innovation. *International Finance Research*, (10), 78–87. <https://doi.org/10.16475/j.cnki.1006-1029.2021.10.008>
- [34] Li, D., & Bao, Y. (2024). The impact of new quality productive forces on firm performance. *Finance and Economics*, 39(5), 40–48.
- [35] Zhang, Z., Zhang, N., & Xiao, T. (2025). Digital and green transformation and firm performance. *Management Review*, 37(11), 67–80. <https://doi.org/10.14120/j.cnki.cn11-5057/f.2025.11.008>
- [36] Tao, X., Chen, Y., Li, D., & Feng, X. (2025). ESG performance, green innovation, and firm performance. *Science & Technology Progress and Policy*, 42(15), 87–97.
- [37] Jiang, T. (2022). Mediation and moderation effects in causal inference research. *China Industrial Economics*, (5), 100–120. <https://doi.org/10.19581/j.cnki.ciejournal.2022.05.005>
- [38] Li, H. (2026). New quality productive forces and rural revitalization. *Statistics & Decision*, 42(1), 11–16. <https://doi.org/10.13546/j.cnki.tjyj.2026.01.002>
- [39] Su, T., Guo, X., Yu, Y., & Wang, Q. (n.d.). Can digital transformation promote collaborative innovation? *Journal of Systems Management*, 1–18.
- [40] Guo, W., & Kang, W. (2024). ESG performance, digital finance, and green technological innovation. *Finance Monthly*, 45(17), 49–54. <https://doi.org/10.19641/j.cnki.42-1290/f.2024.17.009>

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Valuation Analysis of Alibaba Using P/E, P/B, and EV/EBITDA Models

Kairong Yang*

School of International Programs, Guangdong University of Finance, Guangzhou, China

**Corresponding author: Kairong Yang.*

Abstract

Background: As a leading enterprise in China's internet industry, Alibaba Group Holding Limited holds significant reference value for investors, competitors, and academic researchers in terms of valuation analysis. This study aims to evaluate its intrinsic value and assess the rationality of market pricing through multiple relative valuation models. **Methods:** This study selects Tencent Holdings, JD.com, and Meituan as comparable companies. It applies three trading multiples within the relative valuation framework—Price-to-Earnings (P/E), Price-to-Book (P/B), and Enterprise Value to EBITDA (EV/EBITDA)—to conduct a comprehensive valuation analysis of Alibaba for fiscal year 2025 and the forecast period. **Results:** The analysis indicates that among the three models, EV/EBITDA demonstrates the highest applicability for valuing Alibaba in fiscal year 2025, as it is not affected by capital structure, non-operating income, or substantial capital expenditures. The P/E ratio exhibits reduced reliability due to distortions in net profit, while the P/B ratio shows the lowest applicability because Alibaba's core value lies in its asset-light structure and platform ecosystem, which are not well aligned with the assumptions of this model. **Conclusion:** The findings emphasize that the selection of valuation methods should be based on a firm's specific operating conditions and financial characteristics at a given stage.

Keywords

Alibaba, valuation analysis, P/B, EV/EBITDA, comparable company analysis

1. Introduction

Since the late 20th century, the internet industry has rapidly entered public view, leading to the accelerated development of global internet applications and the emergence of a large number of e-commerce companies. According to relevant data and statistical surveys, the number of global internet users has reached approximately 2.4 billion to date, and the internet industry contributes 8.3% and 7.3% to the GDP of countries such as the United Kingdom and South Korea, respectively, demonstrating its substantial growth potential and influence [1]. Amid the rapid expansion of the global internet, China's internet market also experienced significant development. In 1999, with the widespread global adoption of the internet, Jack Ma and 17 co-founders established Alibaba Group Holding Limited to meet the growing demand for e-commerce in China. Within three years of its establishment, Alibaba achieved profitability in 2002 and continued to expand through organic growth and mergers and acquisitions. In 2009, it entered the cloud computing sector and began developing artificial intelligence technologies to support its platform ecosystem.

Over two decades, Alibaba expanded its business through diversification, increasing its revenue to USD 72 billion [2], with operations spanning consulting, e-commerce, finance, technology, and artificial intelligence. With the vision and leadership of Jack Ma, Alibaba rapidly grew into China's largest online commerce enterprise. By 2020, it had become the sixth most valuable company in the world, with 757 million annual active customers [2]. As a result, it has become a favored target for investors and a focal point for valuation analysis among industry analysts. Therefore, conducting a comprehensive and in-depth valuation analysis of Alibaba is of great significance.

The analytical method used to estimate a company's current or projected value is referred to as corporate valuation. Through valuation analysis, a firm's fair value and intrinsic value can be approximated and subsequently compared with its current market value [3]. The core objective of valuation is to determine whether a company's stock is "overvalued" or "undervalued." For investors, in-depth financial modeling can help identify discrepancies between market price and intrinsic value, thereby informing decisions to "buy," "hold," or "sell," uncovering investment opportunities while mitigating risks. For competitors, as an industry leader, Alibaba's financial indicators—such as revenue growth rate—as well as its business model and market strategies serve as industry benchmarks. By analyzing Alibaba's intrinsic value, competitors can better understand industry standards and position themselves strategically within the market. In summary, a comprehensive valuation analysis of Alibaba provides essential insights for decision-making, competition, and academic research, with implications extending beyond the firm itself to the broader industry. Therefore, to assess Alibaba's intrinsic value, this study employs relevant financial data and industry reports, and applies three trading multiples—P/E, P/B, and EV/EBITDA—within the framework of relative valuation to conduct a comprehensive analysis of the same subject: Alibaba.

2. Overview of Valuation Methods

Valuation is a core tool in finance, investment, and corporate management. It is not merely a numerical outcome, but also a framework for understanding value, assessing risk, and making rational decisions. For internet companies, which have attracted extensive attention from equity investors, reasonable valuation serves as a prerequisite for scientific pricing and investment decisions [4]. Therefore, the selection of appropriate valuation methods for internet firms is particularly critical. Relative valuation determines value by benchmarking against comparable companies in the market, whereas absolute valuation estimates intrinsic value based on a firm's expected future cash flows. The combination of these two approaches allows for mutual verification and enhances the robustness of valuation conclusions. Accordingly, this study employs both relative and absolute valuation approaches, utilizing different parameters and models to assess the intrinsic value of Alibaba Group Holding Limited.

Within the relative valuation framework, one of the most widely used core methods is Comparable Company Analysis (Comps). Its essence lies in the identification of comparability or similarity [5], and it serves as a primary approach for valuing specific companies, divisions, business segments, or asset portfolios. The application of Comparable Company Analysis is based on the premise that similar companies share key operational, financial, and risk characteristics, thereby providing meaningful benchmarks for valuation. By positioning the target company relative to its industry peers, analysts can derive appropriate valuation parameters. The core of this approach lies in selecting a group of sufficiently comparable firms (i.e., the "peer group") as references. Subsequently, valuation multiples (trading multiples) are calculated, which form the basis for determining the valuation range of the target company. This range is obtained by applying the selected multiples to the relevant financial metrics of the target firm.

Commonly used valuation multiples include P/E, EV/EBITDA, and P/B, among which the first two are the most frequently applied. The Price-to-Earnings (P/E) ratio is based on equity price and earnings, calculated as the current share price divided by diluted earnings per share (or equity value divided by net income). It can be interpreted as a measure of how much investors are willing to pay for each unit of a company's current or future earnings. As one of the most widely recognized trading multiples, the P/E ratio is typically based on forward earnings per share (and to a lesser extent on trailing twelve-month earnings), reflecting investors' focus on future growth. The formulas are as follows: $P/E \text{ Ratio} = \text{Share Price} / \text{Diluted EPS}$; $P/E \text{ Ratio} = \text{Equity Value} / \text{Net Income}$. Despite its widespread use, the P/E ratio has limitations. For companies with minimal or negative earnings, the denominator may be extremely small or even negative,

rendering the metric unreliable or inapplicable. The Price-to-Book (P/B) ratio is a financial valuation metric used to assess the relationship between a company's market value and its book value. It is calculated as: $P/B = \text{Market Capitalization} / \text{Net Book Value}$; or $P/B = \text{Share Price} / \text{Net Book Value per Share}$, where: $\text{Net Book Value} = \text{Total Assets} - \text{Total Liabilities}$. The Enterprise Value to EBITDA (EV/EBITDA) ratio is a widely adopted valuation standard across most industries. Its advantages lie in its independence from capital structure and taxation, as well as its resistance to distortions arising from differences in accounts receivable and payable across firms. However, it may overestimate cash flow for capital-intensive companies when capital expenditures are not considered, and it does not reflect changes in working capital, which may lead to distortions in actual cash flow. The formula is: $\text{EV/EBITDA Ratio} = \text{Enterprise Value} / \text{EBITDA}$, where: $\text{Enterprise Value (EV)} = \text{Market Capitalization} + \text{Interest-bearing Debt} - \text{Cash Equivalents}$; $\text{EBITDA} = \text{Net Income} + \text{Income Tax} + \text{Interest Expense} + \text{Depreciation and Amortization}$.

3. Selection and Analysis of Comparable Companies

When applying the Comparable Company Analysis method for relative valuation, identifying a set of core comparable firms and comparing their relevant financial and operational data with those of the target company is a crucial step in determining the valuation range. To identify firms with similar business and financial characteristics, a clear and comprehensive understanding of the target company is essential.

Based on research and analysis, this study selects Tencent Holdings, JD.com, and Meituan as comparable companies to Alibaba Group Holding Limited. The rationale for this selection is as follows. First, from the perspective of business operations and industry classification, Alibaba, as the largest e-commerce platform in China, operates across multiple sectors, including e-commerce, cloud computing, logistics, digital media, and local services, forming a comprehensive and robust business ecosystem. Tencent, as an integrated internet service provider, has become a global leader in internet technology, with business segments covering social media, digital entertainment, fintech, and cloud computing [6]. JD.com, as a leading self-operated e-commerce enterprise, operates one of the world's largest drone delivery systems and infrastructures, and is listed among the *Fortune Global 500* companies. Its business model is similar to that of Amazon [7], centering on a hybrid model combining self-operated retail and a marketplace platform, supported by logistics and fintech services, with a strong emphasis on supply chain capabilities. Meituan, as a leading local services platform, primarily operates online and on-demand delivery services. In addition to food delivery, its services have increasingly expanded to include instant consumption goods such as pharmaceuticals and daily necessities [8]. These four companies are all Chinese internet giants listed in the Hong Kong market, with highly related and complementary business operations. Their industries and business models exhibit a high degree of similarity. Second, from the perspective of products and services, Alibaba operates platforms such as Taobao, Tmall, and Alibaba Cloud, adopting a business model that integrates platform operations with technology services. Tencent provides products such as WeChat and QQ, with a service model centered on social connectivity combined with technological support. JD.com operates JD Mall and JD Logistics, following a model that combines self-operated retail with an open platform. Meituan provides services such as Meituan Food Delivery and Meituan Grocery, adopting an O2O (online-to-offline) model integrated with on-demand delivery. All four companies offer online transaction services to both consumers and merchants, and have established a dual-driven model based on technology and services, indicating strong comparability in their product and service structures. Third, from the perspective of customer base and end markets, Alibaba serves users across all age groups, with a primary focus on middle-to-high-income consumers, while also covering multiple industries such as manufacturing and retail. Tencent also serves a broad user base across all age groups, leveraging its strong social traffic advantage, with major clients including game developers and content creators. JD.com primarily targets middle-class consumers in first- and second-tier cities, characterized by strong purchasing power, high repurchase rates, and higher average order values, while serving authorized distributors and retailers with an emphasis on product quality and authenticity assurance. Meituan primarily targets younger consumers, characterized by high-frequency and low average transaction value consumption patterns. Overall, all four companies possess massive user bases and rank among the most widely used internet platforms in China. Their customer segments and end markets are therefore broadly comparable. Finally, the financial conditions of these companies are analyzed across five dimensions, as presented in Table 1.

Table 1: Financial Summary Table

Company	Market Capitalization (Nov 2025)	Revenue (FY2025)	Total Assets (Q2 2025)	Net Profit Margin (FY2025)	Gross Margin (Q2 2025)	Revenue Growth (FY2025)	Capital Expenditure (Q2 2025)	ROE (Q2 2025)	ROA (Q2 2025)	Debt-to-Asset Ratio (Q2 2025)	Cash Reserves (Q2 2025)
Alibaba Group Holding Limited	HKD 3.02 trillion	RMB 996.35 billion (+6%)	RMB 1.2 trillion	14.1%	42%	6%	RMB 31.8 billion	12%	8%	48%	RMB 585.66 billion
Tencent Holdings	HKD 5.95 trillion	RMB 660.26 billion (+10%)	RMB 2.01 trillion	29.7%	56%	10%	RMB 20.0 billion	8.8%	5.7%	40.3%	RMB 500.0 billion
JD.com	HKD 387.9 billion	RMB 1,158.82 billion (+6.8%)	3940 billion	3.85%	15%	6.8%	RMB 15.0 billion	6%	3%	53.5%	RMB 116.5 billion
Meituan	HKD 645.0 billion	RMB 337.59 billion (+22%)	RMB 330.2 billion	10.6%	35%	22%	RMB 12.0 billion	5.7%	3.2%	44.2%	RMB 100.0 billion

Data source: Company financial reports, industry reports, and Sina Finance.

Based on the market capitalization, revenue, and total asset data presented in Table 1, the relative scale of the companies can be assessed. The results indicate that all firms are comparable in size and belong to the top tier of China's internet industry. From the revenue growth rates, it can be observed that although each company is driven by different growth factors, all are currently undergoing business transformation and diversified expansion, suggesting a high degree of comparability. Furthermore, based on indicators such as ROE, ROA, debt-to-asset ratio, and cash reserves, the investment return and financial stability of these firms can be evaluated. The results show that all companies maintain sound financial structures and sufficient cash reserves to withstand risks, thereby providing a solid foundation for relative valuation analysis.

4. Valuation Analysis

This study applies three commonly used trading multiples within the relative valuation framework—P/E, P/B, and EV/EBITDA—to conduct a valuation analysis of Alibaba Group Holding Limited. The valuation date is set as March 2025, and 2025E and 2026E refer to forecast values based on calendar years.

Table 2: Valuation Implied by P/E

Time Period	Financial Metric (billions)	Multiple Range	Implied Equity Value (HKD billions)	Fully Diluted Shares (billions)	Implied Share Price (HKD)
LTM	134.79	14.1x—21.5x	1890.54—2907.99	19.1	98.98—152.25
2025E	141.24	16.0x—23.0x	2259.84—3248.52	19.1	118.32—170.08
2026E	190.46	18.0x—25.0x	3428.28—4761.50	19.1	179.49—249.29

Data source: Company financial reports, announcements from the Hong Kong Monetary Authority, and industry research reports.

Table 3: Valuation Implied by EV/EBITDA

Time Period	Financial Metric (billions)	Multiple Range	Implied Enterprise Value (billions)	Net Debt (billions)	Implied Equity Value (billions)	Fully Diluted Shares (billions)	Implied Share Price (HKD)
LTM	235.4	8.0x—11.0x	1883.2—2589.4	-264.29	2147.49—2853.69	19.1	112.43—149.41
2025E	251.45	9.0x—12.0x	2263.05—3017.40	-264.29	2527.34—3281.69	19.1	132.32—171.82
2026E	299.6	9.5x—12.5x	2846.20—3745.00	-264.29	3110.49—4009.29	19.1	162.85—209.91

Data source: Company financial reports, announcements from the Hong Kong Monetary Authority, and industry research reports.

Table 4: Valuation Implied by P/B

Time Period	Book Value (billions)	Multiple Range	Implied Equity Value (billions)	Fully Diluted Shares (billions)	Implied Share Price
LTM	877.4	2.0x—4.0x	1754.8—3509.6	19.1	91.88—183.75
2025E	941.6	2.2x—4.2x	2071.52—3954.72	19.1	108.46—207.05
2026E	1016.5	2.4x—4.4x	2439.6—4472.6	19.1	127.73—234.17

Data source: Company financial reports, announcements from the Hong Kong Monetary Authority, and industry research reports.

According to data from the Hong Kong Exchange, the median share price and equity value during the LTM (Last Twelve Months) period, with March 2025 as the valuation starting point, are HKD 130.92 and HKD 2,500.59 billion, respectively. The corresponding averages are HKD 131.45 and HKD 2,510.69 billion. By comparing the valuation results derived from the three multiples for the LTM period, it is evident that EV/EBITDA demonstrates the highest applicability in valuing Alibaba. The possible reasons are as follows: First, the P/E ratio relies on net profit. However, in fiscal year 2025, Alibaba's net income was significantly affected by non-operating gains, which undermines the reliability of both the mean and median P/E values as indicators of its core operating performance. Second, to advance its "AI + Cloud" strategy, Alibaba's capital expenditure reached RMB 92.484 billion in fiscal year 2025. This led to a substantial reduction in free cash flow. Meanwhile, increased R&D investment, infrastructure depreciation, and amortization expenses further reduced reported accounting profits. However, these expenditures do not affect EBITDA calculations, making EV/EBITDA a more stable and appropriate valuation metric. In contrast, the P/B ratio exhibits the lowest applicability among the three valuation multiples for Alibaba in fiscal year 2025. This is primarily because Alibaba's core value does not lie in its book net assets, making it difficult for P/B to capture its intrinsic worth. Alibaba's competitive advantages are rooted in its platform ecosystem, technological barriers, and user base—intangible, asset-light factors rather than tangible fixed assets. As a result, both the median and mean P/B values provide limited explanatory power for valuation. Moreover, Alibaba's asset structure is fundamentally misaligned with the typical application scenarios of the P/B ratio. P/B is more suitable for industries such as finance and heavy manufacturing, where operations rely heavily on tangible assets. In contrast, Alibaba's core revenues in fiscal year 2025 are primarily derived from platform-based businesses such as Taobao and Tmall (with a profit margin of 43.5%) and technology services such as Alibaba Cloud, which do not depend on significant net asset turnover. Consequently, the applicability of the P/B ratio is naturally the lowest among the three methods.

5. Conclusion

Based on the analysis conducted in this study, and the valuation calculations for Alibaba Group Holding Limited in fiscal year 2025, it is found that, under specific conditions, EV/EBITDA may be the most appropriate trading multiple for valuation. However, in practice, each of the three valuation multiples exhibits its own advantages and limitations. When Alibaba's capital structure—such as debt and cash levels—remains relatively stable, and net profit accurately reflects its core profitability, the P/E ratio is more suitable. This is because P/E directly corresponds to equity value and aligns with the return perspective of equity investors, providing an intuitive measure of valuation per unit of earnings. When Alibaba experiences significant changes in capital structure or when net profit is substantially affected by non-operating factors, EV/EBITDA becomes more appropriate. By excluding the effects of debt, taxation, and depreciation and amortization, EV/EBITDA focuses on the firm's core operating cash flow capacity and better reflects its operational performance. The P/B ratio, which is linked to net assets, is more suitable in scenarios where profitability is temporarily weak but asset quality remains high and a sufficient margin of safety exists. For example, if the internet advertising industry were to experience a downturn in 2026, leading to a decline in Alibaba's advertising revenue and a corresponding decrease in net profit, the P/E ratio might appear artificially elevated due to reduced earnings and thus fail to reflect the firm's true valuation. Therefore, selecting appropriate valuation methods and trading multiples based on a firm's specific operating conditions, capital structure, and financial indicators—such as cash flow—is a critical task that requires careful consideration by analysts.

Finally, this study only conducts valuation analysis based on Alibaba's financial data from 2025 to 2027. Future research will further refine the valuation framework by incorporating additional trading multiples and

extending the analysis across more time periods, in order to better estimate Alibaba's intrinsic value under varying conditions.

References

- [1] Gao, F. (2017). *Research on valuation methods and applications of internet companies* [Master's thesis, East China Normal University].
- [2] Pai, Y., & Kumar, K. G. N. (2021). E-commerce to multinational conglomerate: Journey of Alibaba Group—A case study. *International Journal of Case Studies in Business, IT and Education (IJCSBE)*, 5(1), 25–35.
- [3] Fang, Z. (2023). Research and application of company valuation methods. *BCP Business & Management*, 45, 109–114.
- [4] Zeng, F., Wu, X., Wu, Y., et al. (2018). Preliminary exploration on the importance and influencing factors of company valuation in the internet industry. *Shanghai Energy Conservation*, (10), 796–799. <https://doi.org/10.13770/j.cnki.issn2095-705x.2018.10.007>
- [5] Yuan, W. (2003). Application and discussion of comparable company analysis in enterprise valuation. *Industrial Technology Economy*, (04), 83–84.
- [6] Tao, C. (2025). The financial performance and investment analysis of Tencent: Balancing growth potential and challenges. In *Proceedings of the 2025 5th International Conference on Enterprise Management and Economic Development (ICEMED 2025)* (p. 391). Springer Nature.
- [7] Jiang, X., Shi, Q., & Yang, X. (n.d.). Analysis on business risk of JD.com and Meituan. [Unpublished manuscript / Journal article].
- [8] Huang, J., Peng, W., & Tian, Y. (n.d.). Research on Meituan's valuation based on qualitative and quantitative analysis. [Unpublished manuscript / Journal article].

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

How the Digital Economy Shapes Workers' Employment Choices: Internal Mechanisms and Empirical Evidence on Informal Employment

Yunshan Xie*

School of International Business, Tianjin Foreign Studies University, Tianjin, China

**Corresponding author: Yunshan Xie.*

Abstract

Using data from the Chinese General Social Survey (CGSS) from 2017 to 2023, this study employs a Probit model to examine the impact of digital economy development on informal employment among the labor force. The results indicate that the development of the digital economy significantly increases the probability of engaging in informal employment. Mechanism analysis shows that the digital economy improves individuals' internet usage, which in turn enhances human capital, reduces information barriers and transaction costs in the labor market, and ultimately raises the likelihood of informal employment. Heterogeneity analysis further reveals that the employment effects of the digital economy are not universally inclusive: the impacts are more pronounced among individuals with non-agricultural household registration, higher educational attainment, and the self-employed. Moreover, regions with higher levels of economic development benefit more from the promotion effect of the digital economy on informal employment.

Keywords

digital economy, informal employment, internet use

1. Introduction

The report of the 20th National Congress of the Communist Party of China emphasizes that employment is the most fundamental component of people's well-being. Employment constitutes the cornerstone of social security and serves as a fundamental driver of economic development. However, against the backdrop of deepening population aging, structural transformation of the demographic dividend, and rising operational costs under traditional employment models, many firms have adopted measures such as hiring freezes and layoffs to reduce labor costs, thereby exerting adverse effects on the employment environment. In July 2020, the *Opinions of the General Office of the State Council on Supporting Multi-Channel Flexible Employment* (Guobanfa [2020] No. 27) explicitly proposed fostering new forms of employment, including part-time work and remote home-based work. As a key mechanism for maintaining labor market resilience, informal employment has increasingly become an indispensable means of expanding employment capacity and improving employment quality.

According to the definition of the International Labour Organization (ILO), informal employment mainly refers to flexible forms of work characterized by the absence of formal labor contracts, fixed working hours,

and direct managerial supervision. It typically includes employment without formal contracts, untaxed income, and a lack of social security coverage and employment benefits. With the acceleration of globalization and marketization, the scale of informal employment in China has continued to expand, playing a positive role in absorbing urban labor, increasing household income, and reducing income inequality. It has gradually become one of the major forms of employment in China [1].

The digital economy, supported by a new generation of information technologies, promotes the evolution of the socio-economic system toward digitalization, networking, and intelligence through deep integration with the real economy [2]. With the advancement of a new round of technological revolution and industrial transformation, the scale of China's digital economy has continued to expand. According to the *China Digital Economy Development Report (2024)*, the size of China's digital economy reached 53.9 trillion yuan in 2023, accounting for 42.8% of GDP, up from 32.9% in 2017, and contributing 66.45% to GDP growth. In recent years, under the macro environment of continuous digital economy expansion, informal employment forms such as self-employment and flexible work have developed rapidly [3]. During the COVID-19 pandemic, the digital economy played a crucial role. In particular, its cross-regional and time-independent online advantages effectively mitigated transmission risks associated with physical gatherings, enabling it to recover earlier and remain more active in the labor market, thereby functioning as a key stabilizer of employment [4].

The rapid development of the digital economy not only creates a large number of new jobs but also profoundly reshapes employment structure and job quality. On the one hand, it accelerates industrial upgrading through digital transformation, thereby optimizing the employment structure [5]. On the other hand, it enhances the efficiency of information transmission, expands access to educational resources, and creates favorable conditions for broadening social capital and improving human capital, which in turn enhances overall employment quality [6]. In addition, new forms of employment emerging from the digital economy diversify the employment choices of younger generations in informal sectors, weakening the intergenerational transmission of employment preferences from parents to children. Meanwhile, the rapid development of the digital economy has gradually narrowed the income gap between informal and formal employment [7].

At present, no unified consensus has been reached in the academic literature regarding the definition of informal employment, and findings on the impact of the digital economy remain mixed. Some scholars argue that the digital economy encourages individuals to choose informal employment. Based on CLDS data, Zhao Xinyu and Zhu Rui find that the development of the digital economy significantly increases the probability of entering informal employment only when individuals are proficient in or frequently use the internet [8]. Using data from the China Family Panel Studies (CFPS), He Zongyue and Song Xuguang further confirm that digital economy development promotes informal employment to a certain extent and has a significant job-creation effect in informal sectors [4]. In contrast, other scholars hold the opposite view. Kuhn and Mansour argue that digital technologies establish network bridges between workers and firms, significantly expanding access to formal employment opportunities [9]. Using provincial panel data from 2002 to 2016, Zeng Xiangjin and Luo Yan find that increased internet penetration facilitates unemployed individuals' access to formal job information, thereby promoting the transition from informal to formal employment [10]. Song Lin and He Yang, based on CFPS 2018 data, demonstrate that internet use improves job matching efficiency and expands employment choices, thus increasing the rate of formal employment [11]. Therefore, exploring the mechanisms through which the digital economy influences labor force decisions regarding informal employment, and accurately identifying the intrinsic relationship between the two, is of important practical significance for better leveraging the role of the digital economy in stabilizing employment and safeguarding livelihoods.

Based on the above analysis, this study utilizes data from the Chinese General Social Survey to empirically examine the impact of digital economy development on informal employment, constructing a digital economy index system. Compared with existing studies, the marginal contributions of this paper are threefold. First, it enriches the literature on digital economy development and informal employment. Most existing studies on the employment effects of the digital economy focus on macro-level analysis or the overall labor force perspective [5][12], paying insufficient attention to employment issues in the digital era. This study serves as both an extension of current digital economy research and an empirical response to the role of the digital economy in enhancing social welfare. Second, it supplements the empirical evidence on informal employment under the new economic landscape. Previous studies have mainly examined informal employment from single perspectives such as digital finance or industrial intelligence [13][14]. This paper constructs a digital economy index from two dimensions—internet development and digital inclusive finance—thereby providing new

evidence on how the digital economy promotes informal employment. Third, it reveals the underlying mechanisms through which digital economy development affects informal employment. By examining individuals' internet usage, this study analyzes the micro-level mechanisms of how the digital economy facilitates informal employment. Furthermore, to test whether the positive effects of the digital economy are universally inclusive, this study conducts heterogeneity analyses based on regional development levels, human capital endowments, and household registration status, thereby providing more nuanced empirical evidence on the employment effects of the digital economy.

2. Theoretical Analysis and Hypothesis Development

2.1 The Impact of Digital Economy Development on Informal Employment

The development of the digital economy weakens the rigid constraints of formal employment systems, strengthens firms' preference for hiring informal workers, and increases the probability of informal employment. By breaking spatial and temporal limitations, the digital economy enables firms to overcome physical constraints in local labor supply, allowing for cross-regional and project-based dynamic allocation of labor. It also promotes the de-formalization of employment relationships, reducing the necessity for long-term labor contracts and social insurance contributions, thereby enhancing flexibility in employment decisions [15]. Intensified market competition under the digital economy further incentivizes firms to reduce costs by hiring temporary workers or outsourcing business activities to informal enterprises, which in turn increases the likelihood of workers entering informal employment [16]. Compared with formal employment, the flexibility, diversity, and lower cost of informal employment make it more attractive to firms. Meanwhile, the development of the labor market and the rise of various intermediary institutions provide organizational support for the expansion of informal employment [17]. Yin Zhifeng et al. further confirm that, after digital empowerment, private enterprises are more inclined to adopt informal employment arrangements to reduce costs [18].

The digital economy also fosters new forms of employment, lowers entry barriers for workers, and increases the probability of informal employment. While improving productivity, the digital economy reduces the prices of goods and services, enhances consumers' real income and consumption levels, and thereby generates greater market demand for informal employment [19]. In this context, a large number of flexible jobs—such as express delivery and food delivery services—have emerged based on digital platforms. These jobs attract young workers who were previously employed in factories due to their flexibility, autonomy, and relatively higher income, thus increasing the share of informal employment [20]. In addition, the rapid development of inclusive financial systems enabled by digital technologies has significantly alleviated financing constraints faced by individuals engaging in self-employment or entrepreneurship, thereby increasing the likelihood of choosing informal employment [8]. Compared with the formal sector, informal employment demonstrates stronger labor absorption capacity: under the same level of capital investment, it can generate two to three times as many jobs as the formal sector [21]. Groups such as individuals with rural household registration, non-members of the Communist Party of China, and those with lower educational attainment often face disadvantages due to weaker factor endowments and higher entry barriers to formal employment, which limits their ability to equally benefit from economic growth. Informal employment, characterized by flexibility and low entry thresholds, provides these groups with opportunities to share in the benefits of digital economy development [7].

Furthermore, the digital economy improves the efficiency of resource allocation, expands access to high-quality employment channels, and increases the probability of informal employment. Since income from informal employment is approximately 1.5 times that of formal employment, and given its higher autonomy, lower entry barriers, and lower operating costs, more migrant workers tend to move from formal to informal employment than vice versa. This pathway has gradually evolved into an important channel for self-employed informal workers to accumulate capital and achieve upward mobility [22]. Leveraging the efficient resource-matching capabilities of digital platforms, informal workers in the digital economy context generally earn higher incomes and enjoy greater autonomy in value orientation and lifestyle choices. As a result, such employment forms are more likely to attract young workers who are proficient in internet use and possess higher education and professional skills to actively choose informal employment [23].

Hypothesis 1: The development of the digital economy increases the probability of informal employment among workers.

2.2 Mechanisms Through Which the Digital Economy Increases Informal Employment

The development of the digital economy relies fundamentally on internet technologies, and the realization of its employment effects depends critically on workers' access to the internet [24]. Individuals who use the internet can overcome spatial and temporal constraints, efficiently obtain employment information, reduce job search costs, accelerate job matching, and increase employment opportunities. Internet use plays a partial mediating role in rural labor participation in informal employment, primarily through enhancing human capital [25]. Through online platforms, rural workers can independently acquire new knowledge and skills, thereby improving their human capital and strengthening their competitiveness in non-agricultural employment [26]. Only for groups that are proficient in or frequently exposed to the internet does the digital economy exert a more significant positive effect on informal employment, which further increases their likelihood of choosing informal employment [8].

Hypothesis 2: The development of the digital economy increases internet use among workers, thereby raising the probability of informal employment.

3. Research Design

3.1 Data Sources

The micro-level data used in this study are drawn from the Chinese General Social Survey (CGSS). Initiated in 2003, the CGSS is the first nationwide, comprehensive, and continuous large-scale social survey project in China, covering multiple levels such as society, households, and individuals, with strong sample representativeness. Based on data availability, this study selects four waves of CGSS data from 2017, 2018, 2021, and 2023. The data are processed as follows: individuals who are currently unemployed are excluded; individuals engaged in agriculture are excluded; individuals under the age of 16 are excluded; and unpaid family workers are excluded. Among the retained samples, individuals without employees (self-employed employers), self-employed business owners, employees without formal labor contracts, casual laborers, dispatched workers, paid family workers, and freelancers are classified as informal workers, while the remaining individuals are classified as formal workers.

The macro-level data are obtained from the *China Statistical Yearbook* and the *Peking University Digital Financial Inclusion Index*. Provincial-level digital economy indicators for the years 2017, 2018, 2021, and 2023 are matched with micro-level informal employment data, yielding a final sample of 8,633 observations.

3.2 Variable Definitions

3.2.1 Dependent Variable

The dependent variable is whether an individual is engaged in informal employment (*informal*). Specifically, *informal* = 1 indicates that the individual is in informal employment, while *informal* = 0 indicates formal employment. Informal employment generally refers to workers who are not fully covered by formal institutional frameworks, including labor regulations, social security systems, statistical registration, and tax supervision, typically excluding the agricultural sector. This study primarily follows the definition proposed by Xue Jinjun and Gao Wenshu, while also incorporating the general framework of the International Labour Organization. Individuals categorized as informal workers include: employees or wage earners without formal labor contracts, employers with zero employees, self-employed workers, and family workers [27]. After excluding unemployed individuals and those engaged in agriculture, the following groups are classified as informal employment (assigned a value of 1): self-employed employers without employees, individual business owners, freelancers, gig workers, casual laborers, paid family workers, employees without labor contracts, and contract-based laborers. All other workers are classified as formal employment (assigned a value of 0).

3.2.2 Core Explanatory Variable

Digital economy development level. Drawing on the framework proposed by Liu Jun et al. and considering the availability of city-level data, this study measures the overall level of digital economy development from two dimensions: internet development and digital financial inclusion [28]. The measurement of internet development follows Huang Qunhui et al., using indicators such as internet penetration rate, employment in related industries, output of related industries, and mobile phone penetration rate [29]. Digital financial inclusion is measured using the *Peking University Digital Financial Inclusion Index*, jointly compiled by the Institute of Digital Finance at Peking University and Ant Group. All indicators are standardized to eliminate differences in measurement units, and the objective entropy weighting method is applied to assign weights and construct a composite digital economy development index.

3.2.3 Mediating Variable

Internet use. Following existing studies, this paper uses internet media usage as a proxy variable for internet use, constructed as a dummy variable [26]. Specifically, individuals who report “never” or “only a few times a year or less” using the internet are defined as infrequent users and assigned a value of 0, while those who report using the internet “several times a week” or “daily” are defined as frequent users and assigned a value of 1.

3.2.4 Control Variables

Following prior literature, this study includes control variables at the individual, household, and regional levels [30]. At the individual level, the control variables include: age and the square of age (the square of age is divided by 100); ethnicity (Han = 1, other ethnic groups = 0); education level (coded from 1 to 13, corresponding to no formal education, private tutoring or literacy classes, primary school, junior high school, vocational high school, general high school, secondary vocational school, technical school, adult junior college, regular junior college, adult undergraduate, regular undergraduate, and postgraduate or above); and household registration status (agricultural hukou = 0, non-agricultural hukou = 1). At the household level, the control variable is family economic status, measured by the logarithm of household income. At the regional level, the control variables include: fiscal support (measured by the ratio of general public budget expenditure to GDP); economic development level (measured by the logarithm of GDP per capita); industrial structure upgrading (measured as the weighted sum of the shares of value added by the primary, secondary, and tertiary industries in GDP, with weights of 1, 2, and 3, respectively); and the urban–rural income gap (measured by the ratio of per capita disposable income of urban residents to that of rural residents).

Descriptive statistics for the above variables are reported in Table 1.

Table 1: Descriptive Statistics of Variables

Variable	Observations	Mean	Std. Dev.	Min	Max
Informal employment	8633	0.576	0.494	0.000	1.000
Digital economy index	8633	0.280	0.114	0.140	0.659
Age	8633	48.676	11.792	22.000	93.000
Age squared	8633	25.084	11.992	4.840	86.490
Ethnicity	8633	0.950	0.218	0.000	1.000
Education level	8633	6.340	3.347	1.000	13.000
Hukou status	8633	0.384	0.486	0.000	1.000
Household income	8633	11.196	0.962	2.079	16.118
Fiscal support	8633	0.214	0.075	0.105	0.608
Economic development level	8633	11.193	0.438	10.319	12.239
Urban–rural income gap	8633	2.376	0.313	1.718	3.438
Industrial structure upgrading	8633	2.506	0.412	2.233	7.430

3.3 Model Specification

To empirically examine the impact of digital economy development on informal employment, this study constructs the following baseline regression model:

$$p^i(informal_{hi}=1)=\alpha_0+\alpha_1digit_{ht}+\sum_{j=1}^n\lambda_jX_{hit}+\delta_t+\theta_h+\mu_{hit} \quad (1)$$

where $\text{informal}_{hi}=1$ indicates that individual i in province h at time t is engaged in informal employment. digit_{ht} denotes the level of digital economy development in province h at time t . X_{hit} represents a set of control variables, including individual-, household-, and regional-level characteristics. In addition, year and province fixed effects are included in the model, where δ_t captures year fixed effects and θ_h captures province fixed effects. μ_{hit} is the random error term.

To test the mediating effect, the following model is constructed based on Equation (1):

$$p^i(\text{informal}_{hi}=1)=\alpha_0+\alpha_1\text{digit}_{ht}+\alpha_2\text{internet}_{ht}+\sum_{j=1}^n\lambda_jX_{hit}+\delta_t+\theta_h+\mu_{hit} \quad (2)$$

where internet_{ht} represents internet use. In this study, internet use is measured as a dummy variable based on individuals' usage of internet media.

4. Empirical Analysis

4.1 Baseline Regression

This study employs a Probit model for the baseline regression analysis. Since the coefficients estimated by the Probit model cannot directly represent the marginal effects of explanatory variables on the dependent variable, Margins estimation is conducted to calculate marginal effects. The coefficients reported in the table therefore reflect the marginal effects of each variable on informal employment. For continuous variables, the marginal effects are evaluated at their mean values. Column (1) reports the regression results including only the digital economy development variable. Columns (2), (3), and (4) present the regression results of digital economy development and informal employment after sequentially controlling for individual-, household-, and regional-level characteristics. The results show that although the absolute values of the coefficients vary across specifications, the direction and statistical significance of the core explanatory variable remain unchanged, indicating that the model estimation results are relatively robust. The main patterns are as follows.

As shown in column (4), holding other factors constant, the digital economy has a positive impact on informal employment and is statistically significant at the 1% level. This indicates that the higher the level of digital economy development in a city, the greater the probability that workers engage in informal employment. Therefore, Hypothesis 1 is supported.

Table 2: Baseline Model: Stepwise Regression

	(1)	(2)	(3)	(4)
Digital economy index	0.948**	1.458***	1.543***	1.700***
	(0.434)	(0.377)	(0.420)	(0.465)
Age		-0.003	-0.004	-0.004
		(0.003)	(0.004)	(0.004)
Age squared		0.005	0.005	0.005
		(0.003)	(0.004)	(0.004)
Ethnicity		-0.006	-0.001	-0.002
		(0.020)	(0.021)	(0.021)
Education level		-0.047***	-0.045***	-0.045***
		(0.002)	(0.002)	(0.002)
Hukou status		-0.117***	-0.114***	-0.114***
		(0.014)	(0.014)	(0.014)
Household income			-0.035***	-0.035***
			(0.007)	(0.007)
Fiscal support				0.477
				(0.615)
Economic development level				-0.286**
				(0.124)
Urban-rural income gap				-0.190
				(0.170)
Industrial structure upgrading				0.010
				(0.016)
Province fixed effects	Yes	Yes	Yes	Yes

Year fixed effects	Yes	Yes	Yes	Yes
Observations	9704	9704	8633	8633
Pseudo R ²	0.078	0.222	0.233	0.234

Note: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively. Robust standard errors are reported in parentheses. The same applies to the tables below.

From the regression results of the micro-level individual characteristics, the coefficient of education level is significantly negative at the 1% level, indicating that workers with higher education are less likely to engage in informal employment. The coefficient of the hukou variable is also significantly negative, suggesting that urban residents have a lower probability of participating in informal employment than rural residents. On average, urban workers are about 11 percentage points less likely to engage in informal employment than rural workers.

From the household-level control variables, the coefficient of household income is negative and significant at the 1% level. Compared with workers from lower-income households, those with higher household income tend to place greater emphasis on job quality and social status, and are therefore more likely to choose formal employment.

From the perspective of macro-level regional characteristics, the coefficient of economic development level is negative and significant at the 1% level, indicating that the more economically developed a region is, the lower the probability that workers engage in informal employment.

4.2 Robustness Checks

To ensure the robustness of the regression results, this study conducts a series of robustness checks from multiple perspectives. The corresponding results are reported in Table 3.

Table 3: Robustness Checks

Sample Treatment	(1) Instrumental Variable	(2) Alternative Sample and Winsorization	(3) Logit
Digital economy index	2.808**	1.205*	1.705***
	(1.326)	(0.645)	(0.449)
Individual controls	Controls	Controls	Controls
Household controls	Controls	Controls	Controls
Regional controls	Controls	Controls	Controls
Province fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	6953	6953	8633
Pseudo R ²	.	0.179	0.235

4.2.1 Endogeneity Issues

Potential endogeneity may arise from reverse causality and omitted variables. This study examines the impact of the overall level of urban digital economy development on individuals' informal employment decisions. While the level of digital economy development may influence individual employment choices, the reverse causal mechanism is unlikely to be significant. Therefore, severe reverse causality is not expected in this context. To further mitigate endogeneity concerns arising from omitted variables, an instrumental variable approach is employed. Following Zhao Tao et al. [31], this study constructs an instrument by interacting the number of fixed telephones per 100 people in each province in 1984 with the national information technology service revenue in the previous period. First, the endogeneity test rejects the null hypothesis that the explanatory variable is exogenous. Second, weak instrument tests indicate that the p-values of the AR and Wald statistics are both significant at the 1% level, suggesting that the selected instrument does not suffer from weak instrument problems. After addressing endogeneity using the instrumental variable, the regression results remain consistent with the baseline findings, confirming that digital economy development significantly promotes informal employment.

4.2.2 Alternative Sample and Winsorization

Considering that municipalities directly under the central government (Beijing, Tianjin, Shanghai, and Chongqing) possess greater autonomy in resource allocation, the employment effects of the digital economy in these regions may exhibit particular characteristics. To ensure the generalizability of the results, these four

municipalities are excluded from the baseline sample. In addition, to mitigate the influence of outliers and extreme values, all continuous variables are winsorized at the 1% level on both tails after excluding the four municipalities. The results are reported in column (2) of Table 3. The findings indicate that, even after excluding municipalities and performing winsorization, the development of the digital economy continues to have a significantly positive impact on informal employment, suggesting that the baseline results are robust.

4.2.3 Alternative Empirical Model

Since the dependent variable is binary, the baseline regression adopts a Probit model. To further verify the robustness of the empirical results, a Logit model—also suitable for binary outcomes—is employed. As shown in column (3) of Table 3, the coefficient of the digital economy index remains significant at the 1% level, indicating that digital economy development continues to increase the probability of informal employment. This further supports the baseline findings of this study.

4.3 Mechanism Analysis

Why does the development of the digital economy increase the probability of informal employment? This study analyzes the underlying mechanism from a micro-level perspective. The development of the digital economy enhances individuals' internet use, which in turn improves human capital, reduces information barriers and transaction costs in the labor market, and expands opportunities for informal employment. This section empirically tests this mechanism, with the results reported in Table 4.

Table 4: Mechanism Test

Variables	(1)	(2)	(3)
	Informal Employment	Internet Use	Informal Employment
Digital economy index	0.693*	0.855**	0.693*
	(0.358)	(0.405)	(0.355)
Internet use			0.034**
			(0.016)
Individual controls	Controls	Controls	Controls
Household controls	Controls	Controls	Controls
Regional controls	Controls	Controls	Controls
Province fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	8033	8033	8033
Pseudo R ²	0.220	0.540	0.220

As shown in columns (2) and (3) of Table 4, the development of the digital economy significantly increases internet use among workers, which in turn raises the probability of informal employment. Therefore, Hypothesis 2 is supported. The development of the digital economy relies fundamentally on internet technologies, and the realization of its employment effects depends on workers' access to the internet [24]. The widespread use of the internet has fundamentally transformed traditional modes of social interaction by providing platforms and tools for online communication and network maintenance. It enables individuals to establish and sustain social connections and expand their social networks. For rural laborers, social interactions—such as participation in social activities—serve as important channels for accessing non-agricultural employment opportunities [32]. In traditional labor markets, migrant workers are often employed in temporary positions and experience high job mobility. The development of the digital economy has increased their use of the internet, thereby leveraging its advantages in providing employment information and reducing transaction and financing costs. This not only enhances their access to job opportunities but also increases the likelihood of informal workers entering more stable employment arrangements. Moreover, the application of internet technologies promotes the development of e-commerce, generating a large number of new forms of employment. Compared with formal employment, informal employment generally imposes lower requirements in terms of education and skills, thus providing more employment opportunities for individuals with lower levels of human capital.

4.4 Heterogeneity Analysis

This section examines the heterogeneous effects of digital economy development on informal employment from both macro- and micro-level perspectives, including regional economic development, hukou status, education level, and employment status. The regression results are reported in Table 5.

Table 5: Heterogeneity Analysis

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Low economic development	High economic development	Non-agricultural hukou	Agricultural hukou	High education	Low education	Employees	Self-employed
Digital economy index	0.797	0.867**	1.009**	0.953*	0.874*	0.647	0.803	0.682**
	(0.568)	(0.382)	(0.416)	(0.537)	(0.523)	(0.451)	(0.731)	(0.326)
Individual controls	Controls	Controls	Controls	Controls	Controls	Controls	Controls	Controls
Household controls	Controls	Controls	Controls	Controls	Controls	Controls	Controls	Controls
Regional controls	Controls	Controls	Controls	Controls	Controls	Controls	Controls	Controls
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4310	4523	3382	5451	3259	5610	1823	7031
Pseudo R ²	0.152	0.234	0.175	0.139	0.106	0.075	0.202	0.208

4.4.1 Heterogeneity by Regional Economic Development

Regional disparities in economic development lead to differences in labor market conditions, which may result in heterogeneous effects of the digital economy on informal employment. Accordingly, this study divides the sample into regions with low and high levels of economic development. Columns (1) and (2) of Table 5 report the regression results. The results indicate that the effect of the digital economy on informal employment is not statistically significant in regions with lower levels of economic development, whereas it is significantly positive at the 5% level in more economically developed regions. The digital economy relies heavily on platform ecosystems, logistics networks, consumption density, and digital infrastructure. In less developed regions, weaker digital infrastructure and insufficient local demand limit the transformation of digital economy development into informal employment opportunities, leading to insignificant effects. By contrast, the promotion effect of the digital economy on informal employment is mainly concentrated in economically developed regions where digital platforms are more prevalent and market demand is stronger.

4.4.2 Heterogeneity by Hukou Status

China's long-standing urban-rural dual structure has led to significant disparities in labor market conditions between individuals with agricultural and non-agricultural hukou, resulting in heterogeneous effects of the digital economy on their employment decisions. This study classifies workers into two groups based on hukou status: non-agricultural and agricultural. Columns (3) and (4) of Table 5 present the regression results. The findings show that the digital economy has a significantly positive effect on informal employment among workers with agricultural hukou at the 10% level, and a significantly positive effect among those with non-agricultural hukou at the 5% level. Relatively speaking, individuals with non-agricultural hukou generally enjoy more stable employment pathways, better social security systems, higher levels of education and public services, and greater advantages in information access, social capital accumulation, and participation in market activities. As a result, they are better positioned to identify, utilize, and convert the opportunities brought by the digital economy, leading to more stable and statistically significant effects.

4.4.3 Heterogeneity by Education Level

The impact of the digital economy on informal employment may differ across workers with varying levels of human capital, and its direction is not a priori clear. To examine this, individuals with a high school education or below are classified as low-education workers, while those with education above high school are classified as high-education workers. Columns (5) and (6) of Table 5 report the results. The findings indicate that the effect of the digital economy on informal employment among low-education workers is not statistically

significant, whereas it is significantly positive at the 10% level for high-education workers. This suggests that, in the digital economy era, low-education workers face difficulties entering skill-intensive digital informal jobs, while traditional manual jobs are more susceptible to replacement, resulting in an insignificant effect. In contrast, highly educated workers, equipped with digital skills and stronger learning capabilities, are better able to seize emerging flexible employment opportunities and adapt to new forms and models of employment generated by the digital economy.

4.4.4 Heterogeneity by Employment Status

Employment status reflects labor–capital relations, and the effects of the digital economy may vary accordingly. This study divides the sample into two groups—employees and self-employed individuals—and conducts separate regressions. Columns (7) and (8) of Table 5 present the results. The findings show that digital economy development significantly increases the probability of informal employment among the self-employed. This is because self-employed individuals generally exhibit higher flexibility and organizational adaptability. The advancement of the digital economy reduces institutional transaction costs, stimulates entrepreneurial incentives, and promotes the expansion of small and micro businesses as well as individual entrepreneurs. By contrast, the effect of the digital economy on informal employment among employees is positive but not statistically significant. On the one hand, the rise of digital platforms has indeed created diverse new employment opportunities. On the other hand, it has simultaneously prompted firms to optimize their employment structures, phasing out low-skilled informal positions. As a result, the net effect on employees' participation in informal employment is not statistically significant.

5. Conclusions and Implications

Employment is the most fundamental component of people's well-being. Against the backdrop of the profound impact of digital economy development on the labor market, this study uses data from the Chinese General Social Survey (CGSS) from 2017 to 2023 and employs a Probit model to examine the effect of digital economy development on the probability of informal employment, as well as its underlying mechanisms. The empirical results show that the development of the digital economy significantly increases the probability of informal employment. Specifically, the digital economy promotes informal employment by enhancing individuals' internet use, which in turn increases their likelihood of engaging in informal employment. Furthermore, the effect of the digital economy on informal employment exhibits significant heterogeneity. The impact is more pronounced among workers with non-agricultural hukou, those with higher levels of education, and the self-employed. The positive effect of the digital economy on informal employment is not universally inclusive, and regions with higher levels of economic development benefit more from this effect. Based on the above findings, the following policy implications are proposed:

First, it is essential to improve the regulatory and policy framework supporting informal employment. Efforts should be made to standardize the development of recruitment platforms to ensure dynamic matching between job information and job seekers' needs, thereby reducing information asymmetry between employers and employees. Policymakers should also pay close attention to the diverse forms of flexible employment emerging in the service sector under the rapid development of the digital economy, clarify labor relationships, and regulate the signing of employment contracts. In addition, it is necessary to explore labor protection systems distinct from traditional employment models, continuously improve social security and employment service systems for flexible workers, and provide institutional support for high-quality employment.

Second, employment-friendly digital transformation should be promoted. Governments should actively foster diverse employment forms derived from the new economy and emerging industries, fully leveraging the role of the digital economy in expanding employment opportunities, reducing physical labor intensity, and enhancing workplace safety. Greater efforts should be made to strengthen the application of internet technologies in key areas such as basic services, public services, and commercial transactions, thereby creating a digital ecosystem conducive to workers' development. At the same time, regulatory mechanisms for the digital economy should be improved, utilizing big data analytics and cloud-based platforms to enhance regulatory efficiency.

Third, the accessibility advantages of digital technology in education should be fully utilized. Governments should strengthen targeted support for informal workers, particularly low-education workers who are forced

into informal sectors due to industrial upgrading or limited digital skills. Through re-employment training and personalized career guidance, their knowledge and professional skills can be effectively improved. Moreover, market-oriented training institutions and industry associations should be encouraged to participate in capacity-building initiatives, providing training programs for informal workers and expanding employment choices for low-education labor groups.

Fourth, it is important to enhance the social recognition of informal workers. Psychological support and humanistic care should be strengthened to foster an inclusive, fair, and respectful working environment, reduce feelings of relative deprivation, and improve job satisfaction and a sense of professional achievement. This would help promote both material and psychological well-being. At the same time, the positive role of informal employment in stimulating market vitality should be objectively recognized, as it can facilitate social mobility, alleviate class solidification, and contribute to the development of more harmonious labor relations.

References

- [1] Zhang, Y., & Zhang, L. (2017). Formation mechanisms and heterogeneity of informal employment in China: Applicability of three major theories. *Population Journal*, 39(2), 88–99.
- [2] Bukht, R., & Heeks, R. (2017). Defining, conceptualising and measuring the digital economy. *SSRN Electronic Journal*.
- [3] Huang, L., & Li, C. (2025). Cross-border e-commerce and urban informal employment. *Journal of International Trade Issues*, (3), 76–94.
- [4] He, Z., & Song, X. (2020). Mechanisms and implications of digital economy in promoting employment: Reflections after the COVID-19 outbreak. *Economist*, (5), 58–68.
- [5] Qi, Y., Liu, C., & Ding, S. (2020). Digital economy development, employment structure optimization, and employment quality improvement. *Economic Perspectives*, (11), 17–35.
- [6] Cong, Y., & Yan, M. (2022). Digital economy, human capital investment, and high-quality employment. *Finance & Economics*, (3), 112–122.
- [7] Li, S. (2025). How informal employment affects intergenerational mobility in the digital economy. *Economy and Management*, 39(4), 32–41.
- [8] Zhao, X., & Zhu, R. (2022). Digital economy and informal employment: Evidence from the China Labor-force Dynamics Survey. *Jilin University Journal Social Sciences Edition*, 62(5), 72–83.
- [9] Kuhn, P., & Mansour, H. (2014). Is internet job search still ineffective? *The Economic Journal*, 124(581), 1213–1233.
- [10] Zeng, X., & Luo, Y. (2019). The impact of internet penetration on informal employment. *Journal of China Institute of Industrial Relations*, 33(3), 47–56.
- [11] Song, L., & He, Y. (2020). The impact of internet use on rural labor employment choices in China. *Chinese Journal of Population Science*, (3), 61–74.
- [12] Bai, P., & Zhang, Y. (2021). Digital economy, declining demographic dividend, and rights of low- and middle-skilled workers. *Economic Research Journal*, 56(5), 91–108.
- [13] Guo, Q., Meng, S., & Mao, Y. (2022). Can digital financial inclusion improve employment quality? *Journal of Shanghai University of Finance and Economics*, 24(1), 61–75.
- [14] Qi, L., & Tao, J. (2023). Industrial intelligence and the employment quality of migrant workers: Mechanisms and improvement paths. *Journal of Huazhong Agricultural University (Social Sciences Edition)*, (1), 34–46.
- [15] Li, G. (2023). De-laborization of platform employment in the digital economy era: Characteristics, challenges, and responses. *Social Sciences in Yunnan*, (2), 151–160.
- [16] Lü, K., Chen, S., & Jia, L. (2023). How does the digital economy affect employment choices? Evidence from work motivation mechanisms. *Research on Economics and Management*, 44(12), 24–43.

- [17] Qian, J. (2022). “Formalization” of informality: Labor market segmentation and domestic service enterprises—Evidence from a 2019 survey in four cities. *Collection of Women’s Studies*, (1), 35–51.
- [18] Yin, Z., Li, H., Liu, Z., & Zhou, M. (2025). Digital empowerment and flexible employment in Chinese private enterprises. *Labor Economics Research*, 13(3), 106–138.
- [19] Zhao, X., & Zheng, G. (2020). Labor market distortions and informal employment: Evidence from the China Labor-force Dynamics Survey. *Jilin University Journal Social Sciences Edition*, 60(4), 163–173.
- [20] Tang, X., & Wang, Z. (2025). The impact of urban service industry development on informal employment. *Shanghai Economic Review*, (10), 91–103.
- [21] Zhang, Y. (2010). Informal employment: Conceptual clarification and value considerations. *Nanjing Journal of Social Sciences*, (4), 62–68.
- [22] Wan, X. (2008). Entry conditions and effects of migrant workers’ informal employment. *Management World*, (1), 63–74.
- [23] Qi, Y., Ding, S., & Liu, C. (2021). Internet use and labor supply of flexible workers in the digital economy: Theory and evidence. *Contemporary Finance & Economics*, (5), 3–16.
- [24] He, Z., Zhang, X., & Wan, G. (2020). Digital finance, digital divide, and multidimensional poverty. *Statistical Research*, 37(10), 79–89.
- [25] He, Q., Liu, G., & Zou, X. (2022). The impact of internet use on rural labor employment choices: Evidence from formal and informal employment. *Journal of Agro-Forestry Economics and Management*, 21(4), 385–394.
- [26] Qi, Y., & Chu, X. (2021). Employment effects of digital life: Mechanisms and micro evidence. *Finance & Trade Economics*, 42(4), 98–114.
- [27] Xue, J., & Gao, W. (2012). Informal employment in urban China: Scale, characteristics, and income inequality. *Comparative Economic & Social Systems*, (6), 59–69.
- [28] Liu, J., Yang, Y., & Zhang, S. (2020). Measurement and driving factors of China’s digital economy. *Shanghai Economic Review*, (6), 81–96.
- [29] Huang, Q., Yu, Y., & Zhang, S. (2019). Internet development and productivity improvement in manufacturing: Mechanisms and evidence from China. *China Industrial Economics*, (8), 5–23.
- [30] He, Y., & Liu, Z. (2024). The impact of digital economy development on informal employment. *Economic Development Research*, (2), 77–90.
- [31] Zhao, T., Zhang, Z., & Liang, S. (2020). Digital economy, entrepreneurial activity, and high-quality development: Evidence from Chinese cities. *Management World*, 36(10), 65–76.
- [32] DiMaggio, P., & Bonikowski, B. (2008). Make money surfing the web? The impact of internet use on the earnings of U.S. workers. *American Sociological Review*, 73(2), 227–250.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Balanced Scorecard and Its Application in Enterprises: A Case Study of ANTA Sports Products Co., Ltd.

Jinwei Fang*

School of Accounting, Dongbei University of Finance and Economics, Dalian 116000, China

**Corresponding author: Jinwei Fang.*

Abstract

In an increasingly competitive sportswear industry, where multiple brands are pursuing accelerated globalization, relying solely on financial metrics to evaluate corporate performance no longer suffices to support the effective execution of strategy. The Balanced Scorecard (BSC) integrates financial indicators with non-financial dimensions—such as customer satisfaction—enabling multi-brand conglomerates to more comprehensively balance short-term profitability against long-term sustainable development. This study selects ANTA Sports Products Co., Ltd. as its case, examining the practical implementation of the Balanced Scorecard across its four core perspectives. The analysis reveals that the BSC's multidimensional evaluation framework aligns well with ANTA's prevailing strategic orientation. Nevertheless, opportunities for refinement remain, particularly in the weighting and prioritization of individual performance indicators. The findings and related conclusions may offer valuable insights for other enterprises engaged in the manufacturing and distribution of sportswear as they develop or refine their own performance evaluation systems.

Keywords

balanced scorecard, ANTA Sports, multi-brand strategy, performance management, globalization

1. Introduction

1.1 Research Background

As competition in the sportswear industry intensifies, multi-brand and globalization-oriented business models have gained widespread adoption. This evolution has rendered corporate performance management increasingly complex. Traditional performance evaluation systems, which rely predominantly on financial indicators, fall short in comprehensively capturing the outcomes of strategic implementation. They also struggle to balance short-term results with long-term development objectives.

Within multi-brand conglomerates, individual brands often occupy distinct market positions and developmental stages, creating inherent tensions between the need for standardized indicators and the demand for brand-specific relevance. The Balanced Scorecard (BSC) addresses these challenges by incorporating non-financial perspectives—such as customer, internal processes, and learning and growth—alongside financial measures. This multidimensional approach translates strategic objectives into actionable,

measurable performance indicators, encouraging organizations to prioritize not only immediate financial outcomes but also capability building and strategic alignment.

In recent years, ANTA Sports Products Co., Ltd. has transitioned from a single-brand operation to a multi-brand, globally oriented enterprise. This transformation provides a compelling real-world context for examining the application of the Balanced Scorecard.

1.2 Research Significance

From a theoretical perspective, existing studies on the application of the Balanced Scorecard in the sportswear industry—particularly within the context of multi-brand conglomerates—remain limited. By taking ANTA Sports as a focal case and adopting the distinctive lens of multi-brand synergy combined with globalization strategy, this paper conducts a detailed examination of how the Balanced Scorecard is actually implemented at ANTA. It aims to contribute fresh insights and extend the body of knowledge in this specialized domain.

On the practical level, ANTA represents a leading exemplar of a multi-brand group in China's sportswear sector. Its experience therefore offers valuable lessons. The study systematically documents both the achievements and shortcomings ANTA has encountered in applying the Balanced Scorecard, thereby providing meaningful and actionable reference points for other enterprises operating under similar structural and strategic conditions.

1.3 Research Methodology

This study employs the case study approach, with ANTA Sports Products Co., Ltd. serving as the primary object of investigation. Through the complementary use of literature review, it systematically analyzes the application of the Balanced Scorecard across ANTA's four core perspectives.

The majority of the data and materials utilized in this research are drawn from the company's publicly disclosed annual reports and other openly available information. The analysis does not incorporate internal, proprietary details concerning ANTA's specific performance evaluation mechanisms. Consequently, certain interpretive elements and conclusions presented herein inevitably involve a degree of inference and reasoned speculation.

2. Theoretical Foundation of the Balanced Scorecard

2.1 Core Essence of the Balanced Scorecard

The Balanced Scorecard (BSC) was introduced by Kaplan and Norton in 1992 as a performance management tool designed to overcome the limitations of traditional financial metrics [1]. Its central premise is to translate an organization's strategy into a set of measurable and evaluable performance indicators. The framework constructs a comprehensive performance evaluation system encompassing four interrelated perspectives: financial, customer, internal business processes, and learning and growth. Through this structure, the Balanced Scorecard achieves balance across four critical dimensions: short-term performance versus long-term development, financial versus non-financial indicators, internal versus external focus, and outcomes versus processes.

Compared with conventional financially oriented performance assessment, the Balanced Scorecard not only emphasizes current profitability but also attaches equal importance to the core capabilities that drive sustainable progress—such as research and development, customer relationships, and operational efficiency. By integrating and mutually reinforcing these four perspectives, the BSC enables strategic objectives to be systematically cascaded and effectively executed throughout the organization.

2.2 Application Value of the Balanced Scorecard

For multi-brand, globally oriented enterprises, the value of applying the Balanced Scorecard manifests primarily in three areas.

First, it enhances strategic visibility. Abstract concepts such as multi-brand synergy and globalization expansion are decomposed into concrete, actionable indicators across the four perspectives, providing clear directional guidance to every business unit.

Second, it coordinates diverse performance requirements. The framework accommodates differentiated performance targets across individual brands and markets while ensuring alignment with the group's overarching strategic objectives.

Third, it fosters long-term competitiveness. By incorporating the learning and growth perspective, the Balanced Scorecard institutionalizes sustained investment in areas such as innovation, research and development, and talent development, thereby laying a robust foundation for enduring future success.

3. Case Background of ANTA Sports

3.1 Company Overview

ANTA Sports Products Co., Ltd., commonly referred to as ANTA, was founded in 1991 and is headquartered in Jinjiang, Fujian Province, China. It stands as China's largest integrated sportswear enterprise and ranks as the world's third-largest sportswear company. Over more than three decades of development, ANTA has evolved into a distinctive "single-focus, multi-brand, globalization" business model.

Its portfolio includes the mass-market sportswear brand ANTA, the premium fashion-oriented brand FILA, and a suite of internationally recognized professional outdoor and specialist brands such as Arc'teryx, Descente, and Salomon. The company's product offerings span the full price spectrum—from entry-level to high-end—and cover diverse usage scenarios from indoor to outdoor activities. This multi-brand parallel approach has enabled sustained expansion; however, it has also imposed greater demands on internal management and performance coordination across the group.

3.2 Multi-Brand and Globalization Development Strategy

ANTA has consistently adhered to the strategic pathway of "single focus, multi-brand, globalization." The "single focus" principle reflects a deliberate commitment to concentrating resources and expertise exclusively within the sportswear industry, eschewing diversification into unrelated sectors. The "multi-brand" dimension involves deploying distinct brands to address the needs and preferences of varied consumer segments. Finally, "globalization" entails extending the company's brands, market presence, and resource allocation toward international markets on a broad scale.

While the multi-brand strategy significantly broadens market coverage, it also introduces considerable management complexity. Individual brands differ markedly in target audiences, price positioning, and stage of maturity. At the same time, the group must preserve each brand's operational independence while fostering strategic coherence and enabling resource sharing across the portfolio. This dual requirement—for unified strategic direction combined with sufficient flexibility—directly shapes the design of ANTA's performance evaluation system.

3.3 Current State of Performance Management and Emerging Challenges

As the company has continued to scale and its business structure has grown more intricate, the limitations of ANTA's earlier financially dominated performance assessment approach have become increasingly apparent. An excessive emphasis on revenue and profit metrics fails to fully reflect the genuine contributions made by individual brands in areas such as market expansion, customer value creation, and organizational capability building.

In the context of multi-brand operations and globalization, different business units naturally prioritize distinct developmental focuses and warrant tailored evaluation emphases. Applying a uniform performance standard across the group can diminish motivational effectiveness and potentially undermine the consistency of strategy execution. Consequently, developing a performance management system that remains aligned with the group's overall strategic objectives while accommodating brand-specific differentiation has emerged as a critical challenge that ANTA must address in its ongoing operations.

4. Practical Application of the Balanced Scorecard at ANTA Sports

ANTA has leveraged the four-perspective model of the Balanced Scorecard to construct a two-tier performance evaluation system tailored to its multi-brand and internationalized characteristics: “unified headquarters-level dimensions + differentiated brand-level dimensions.” At the headquarters level, the group formulates overall strategic plans and decomposes them into key indicators across the four perspectives. Each brand then adapts and further refines these dimensions in line with its own positioning and developmental stage, ensuring that performance indicators remain aligned with group-wide strategic objectives while accommodating the unique growth trajectories of individual brands. The specific implementation is outlined as follows.

4.1 Financial Perspective: Balancing Scale Expansion with Profit Quality

The financial perspective focuses on profitability and resource utilization efficiency, serving as the most direct reflection of strategic outcomes within the Balanced Scorecard framework. At ANTA, the financial dimension emphasizes healthy, high-quality growth: while pursuing overall revenue expansion across segmented and multi-market operations, the company simultaneously prioritizes the profitability of each individual brand.

In practice, the group headquarters establishes core indicators such as revenue growth rate, net profit margin, asset-liability ratio, and cash flow performance. Individual brands then develop more granular financial targets beneath this umbrella, resulting in differentiated evaluation standards. The core ANTA brand concentrates on market share expansion and revenue growth; FILA places primary emphasis on premium profitability and margin improvement; emerging professional brands prioritize growth potential; and overseas brands focus on contributions to localized revenue generation. In implementation, the financial perspective has supported steady overall performance improvement across the group while providing each brand with clear profitability guidance. This dual effect fosters stability at the corporate level and ensures distinct, motivating objectives for every business unit.

4.2 Customer Perspective: Emphasizing Segmented Market Penetration and Customer Value Enhancement

The customer perspective centers on competitive positioning within target markets, together with customer satisfaction and loyalty. ANTA designs differentiated indicators for each brand’s distinct customer segments as follows.

The core ANTA brand targets the “mass professional market.” Key evaluation metrics include market share, membership penetration rate, and repeat-purchase performance of signature high-tech products (such as those featuring Nitrogen Technology), reinforcing the brand’s image of superior value-for-money and widespread technological accessibility.

FILA positions itself in the “fashion-sport market” and prioritizes indicators such as premium-channel sales efficiency (e.g., store productivity per square meter), customer experience satisfaction scores, and overall brand equity strength—measures that help sustain its leadership in fashion-forward sportswear.

Specialist brands such as Descente and Kolon target the “high-end professional market.” Their performance is assessed through metrics including the proportion of stores located in upscale retail environments, engagement levels within core user communities, and sales growth in specialized product categories, thereby enabling deeper penetration into premium professional segments.

For overseas markets, the focus shifts to revenue growth rates in new territories and the effectiveness of integration achieved through group-level channel capabilities. This approach ensures that each brand clearly identifies and serves its intended customer groups within shared competitive arenas.

4.3 Internal Process Perspective: Optimizing Supply Chain Efficiency and Multi-Brand Collaboration

The internal process perspective emphasizes the speed and effectiveness of critical business processes, encompassing supply chain management, research and development, and cross-brand collaboration. ANTA

has implemented a unified supply chain control platform to monitor inventory turnover and order delivery timelines while promoting the sharing of R&D resources and shortening new product development cycles.

First, in terms of supply chain operations, the establishment of a shared operations platform enables centralized monitoring and continuous improvement of inventory turnover days and order fulfillment lead times across the entire group. This facilitates flexible resource allocation and rapid response capabilities.

Second, regarding R&D and innovation, ANTA has developed a centralized R&D platform that enables key material technologies—such as Nitrogen Technology and Cool Skin Technology—to be shared across brands while allowing for brand-specific adaptations. This approach has substantially reduced duplicated R&D expenditures and significantly shortened the time required to translate technological advancements into marketable products. For example, advanced technologies acquired through integrations, such as Texapore waterproofing, have been systematically incorporated into outdoor product lines, thereby enhancing professional-grade performance.

In parallel, ANTA has continuously advanced the integration and standardization of logistics systems, IT infrastructure, and retail operations knowledge across multiple brands. By incorporating indicators such as “resource sharing rate,” the company has achieved tangible reductions in operating costs and realized economies of scale. This perspective has played a particularly positive role in addressing collaboration challenges among brands and has markedly improved overall operational efficiency and process integration.

4.4 Learning and Growth Perspective: Strengthening R&D Innovation, Talent Development, and Digital Capabilities

The learning and growth perspective focuses on talent cultivation, technological innovation, and organizational capability building. Organizational learning and growth constitute the foundational driver for achieving strategic objectives and serve as the core engine for product and technological innovation, manufacturing process improvement, and marketing model evolution [2].

ANTA enhances its long-term strategic capabilities through sustained investment in R&D, development of core technologies, global talent cultivation, and digital transformation. Specific initiatives include:

- The deployment of the “ANTA Academy” digital learning platform and specialized “4 Forces” training courses covering merchandising, retail, and design disciplines, enabling large-scale, precision-targeted employee empowerment programs. The group allocates substantial annual expenditures to training activities.
- Vigorous promotion of digital transformation, including the construction of a comprehensive omni-channel data middleware platform that improves data-supported decision-making ratios and process automation levels across manufacturing, distribution, and sales stages.

Most notably, ANTA has integrated Environmental, Social, and Governance (ESG) objectives into its overarching strategic framework. The company has established a board-level Sustainability Committee and incorporated quantifiable ESG-related key performance indicators (KPIs) into the annual performance evaluations of business unit leaders. This ensures the achievement of targets such as increasing the proportion of sustainable products (reaching over 30% in 2024, for example). These developments indicate that ANTA’s Balanced Scorecard has evolved into an advanced framework that incorporates a broader set of long-term value creation factors.

5. Evaluation of the Application Effectiveness of the Balanced Scorecard at Anta Sports

5.1 Steady Improvement in Financial Performance and Optimization of the Multi-Brand Profit Structure

Under the Balanced Scorecard framework, Anta Sports has successfully achieved the dual objectives of robust financial performance growth and diversification of its profit structure. From 2020 to 2024, the company’s operating revenue increased from RMB 35.51 billion to RMB 70.826 billion, with a compound annual growth rate (CAGR) of 18.8%. Net profit rose from RMB 5.56 billion to RMB 11.307 billion,

registering a CAGR of 19.2%. The debt-to-asset ratio was consistently maintained below 45%, keeping financial risk within a controllable range.

In terms of profit structure, the proportion of sales revenue from newly acquired professional brands in total operating revenue increased from 8.5% in 2020 to 18.2% in 2024, while the share of revenue from international brands rose from 12.3% to 22.5%. These figures indicate that Anta Sports has reduced its reliance on its core domestic brand and market, achieving a more diversified and resilient profit structure. This demonstrates that the Balanced Scorecard has effectively realized its goals of enhancing short-term profitability while promoting long-term profit diversification.

5.2 Enhanced Competitiveness in Segmented Markets and Continuous Improvement in Customer Value

The differentiated performance evaluation in the customer dimension of the Balanced Scorecard has strengthened the competitive capabilities of Anta's various brands within their respective niche markets. In 2024, Anta's main brand maintained a leading market share in the mass market, FILA held a dominant position in the premium segment, and Arc'teryx led in the professional outdoor market. This has created a comprehensive competitive advantage covering all major market segments.

Furthermore, customer satisfaction and loyalty metrics across all brands have shown year-on-year improvement. Notable indicators include the repurchase rate of Anta's main brand through e-commerce channels and the core user loyalty rate of Arc'teryx, both of which performed exceptionally well. Customer value has been continuously optimized. Guided by the key performance indicators in the customer dimension, each brand has been able to precisely identify and effectively meet the needs of its target consumer groups. As a result, product reputation and brand awareness have improved naturally, laying a solid foundation of loyal customers for future sustainable development.

5.3 Improved Internal Operational Efficiency and Prominent Synergistic Effects among Multi-Brands

Continuous optimization of internal processes has enhanced the overall operational efficiency of Anta Sports Group, with synergistic effects among different brands becoming increasingly evident. According to 2024 data, the group's inventory turnover days decreased by 15 days compared with the previous year, and the average order delivery time continued to shorten. The efficiency of the entire supply chain remains at the forefront of the industry.

Moreover, by sharing and collaborating on supply chain and R&D resources across multiple brands, the group has significantly reduced overall operating costs and accelerated the speed of new product launches. By integrating the channel networks of international brands such as Salomon, these brands have achieved rapid market expansion in their target regions. From a holistic perspective, the improvement in internal operational efficiency and the refinement of coordination mechanisms have provided strong support for the effective implementation of the group's overall strategy.

5.4 Cultivation of Long-Term Development Capabilities and Continuous Enhancement of Core Competitiveness

The learning and growth dimension of the Balanced Scorecard has effectively fostered Anta Sports' long-term development capabilities. Performance indicators in this dimension show that the company has continuously increased R&D investment, achieved substantial improvements in key technologies, and optimized both the retention rate of core talent and the proportion of global talent. Following the implementation of initiatives such as digital transformation, the company's products have gained stronger core competitiveness. The talent pool has become better aligned with the group's globalization strategy, and efficiency across the entire value chain has improved.

These developments have enabled Anta Sports to establish significant advantages in technology, talent, and digital capabilities amid fierce industry competition, thereby providing a solid foundation for realizing its ambitious vision of becoming a global leader in the sporting goods industry.

6. Problems Existing in the Application of the Balanced Scorecard at Anta Sports

6.1 Unreasonable Weight Allocation of Multi-Brand Indicators and Insufficient Synergy

Although Anta Sports has established a differentiated indicator system across its multi-brand portfolio, the allocation of indicator weights remains unreasonable in several aspects. On the one hand, certain brands place excessive emphasis on financial indicators. For example, the profit margin indicator accounts for 40% of FILA's performance evaluation. This has resulted in insufficient investment in non-financial aspects such as brand building and customer experience. In 2024, FILA's marketing expenditure declined by 5% compared with the previous year, and the progress of building its premium brand image lagged behind expectations. Overemphasizing financial indicators in the application of the Balanced Scorecard leads to serious limitations and lagging effects in the performance evaluation system, preventing it from comprehensively and prospectively reflecting the enterprise's actual operations and development trajectory [3].

On the other hand, the weights assigned to synergy-related indicators at the group level are relatively low. For instance, the indicator for multi-brand resource sharing rate accounts for only 10%. This has caused some brands to operate in isolation ("each fighting its own battle"), preventing the full realization of synergistic effects in areas such as supply chain resource sharing and channel integration. Taking the Anta main brand and Salomon as examples, their supply chain resource sharing rate stands at only 35%, which is far below the group's overall average.

6.2 Insufficient Adaptability of Globalization Indicators and Inadequate Consideration of Regional Differences

In the process of implementing its globalization strategy, Anta Sports has not sufficiently accounted for the differences among various regional markets, resulting in the poor suitability of certain globalization indicators. For example, the Salomon brand in the European market continues to use the same "revenue growth rate" indicator (weighted at 35%) as domestic brands. However, the outdoor sports market in Europe is highly competitive with limited growth potential. Setting this indicator too high has caused the European team to overemphasize short-term revenue performance while neglecting local product research and development as well as customer relationship maintenance.

Furthermore, the company lacks region-specific indicators tailored to local policies, consumption habits, and other contextual factors (such as the proportion of localized products and regional policy alignment scores). As a result, the localized operations of some overseas brands have been suboptimal. In the North American market, for instance, the in-store customer satisfaction score for Anta's main brand is only 78 points, which is even lower than in the domestic market.

6.3 Difficulty in Quantifying Non-Financial Indicators and Low Efficiency of Data Collection

Non-financial indicators in the customer and learning & growth dimensions of the Balanced Scorecard are inherently difficult to quantify, and data collection is relatively slow. Taking the "brand reputation" indicator in the customer dimension as an example, Anta relies on data from third-party survey companies, with the entire survey process taking up to three months. Consequently, data updates are not timely. In the learning & growth dimension, the indicator "employee innovation capability" lacks clear quantitative criteria and is indirectly measured only by the number of patent applications, which fails to comprehensively reflect employees' actual innovation level.

In addition, non-financial data across many brands and regions remain fragmented, leading to low efficiency in data aggregation at the group level. As a result, feedback from the evaluation of non-financial indicators is delayed and cannot provide immediate guidance for operational improvement.

6.4 Single Application of Performance Results and Insufficient Incentive Orientation

The performance outcomes of Anta Sports' Balanced Scorecard are primarily applied to employee compensation evaluation, resulting in a rather narrow scope of application and limited motivational effect.

First, performance results have little connection with employee training, promotion, or career development, causing employees to focus solely on meeting performance targets while neglecting the improvement of their own capabilities. Second, insufficient emphasis has been placed on long-term strategic indicators that require extended periods to yield results, such as multi-brand collaboration and global operations. For example, the group-level indicator “increase in the proportion of overseas market revenue” is not included in the main reward scheme. Consequently, individual brands show low motivation for overseas market expansion, and it is projected that by the end of 2024, the overseas market development progress of some brands will fall behind the scheduled plan.

7. Optimization Suggestions for the Application of the Balanced Scorecard at Anta Sports

7.1 Optimize Multi-Brand Indicator Weights and Strengthen Synergy Orientation

To address the issue of unreasonable weight allocation across multi-brand indicators, Anta Sports should refine the indicator weights according to each brand’s strategic positioning, with a stronger emphasis on synergy orientation.

First, the weights of financial and non-financial indicators for individual brands should be adjusted. For FILA, the weight of the profit margin indicator should be reduced from 40% to 25%, while the weights of indicators related to brand building and customer experience should be increased. This adjustment will ensure sustained investment in long-term value creation for the premium brand.

Second, the weights of group-level synergy indicators should be enhanced. For instance, indicators such as the multi-brand resource sharing rate and channel integration speed should be raised to 20% and incorporated into the key performance evaluation criteria of each brand. Brands that demonstrate outstanding performance in synergy outcomes should receive additional incentives. These measures will effectively promote resource integration and the realization of synergistic effects.

7.2 Refine Globalization Indicators with Differentiation and Align with Regional Market Needs

To resolve the mismatch of globalization indicators, Anta Sports should develop differentiated indicators tailored to the characteristics of each regional market.

First, region-specific core indicators should be established. For example, for the Salomon brand in the European market, the weight of the revenue growth rate indicator may be appropriately reduced, while greater emphasis should be placed on indicators such as the proportion of localized products and regional customer satisfaction. For Anta’s main brand in North America, new indicators such as “localized marketing effectiveness” and “regional policy adaptability” should be introduced to guide localized operations.

Second, a dynamic adjustment mechanism for regional market indicators should be established. Indicator settings should be regularly reviewed and modified in response to changes in market competition dynamics and policy environments, thereby improving the overall relevance and adaptability of the performance measurement system.

7.3 Improve the Quantification System for Non-Financial Indicators and Enhance Data Collection Efficiency

To overcome the challenges of quantifying non-financial indicators and the slow pace of data collection, Anta Sports needs to develop a comprehensive quantification system for non-financial indicators and improve the efficiency of data collection and integration.

First, standardized quantification criteria for non-financial indicators should be formulated. For instance, the “brand reputation” indicator can be linked to real-time customer reviews on e-commerce platforms and public sentiment dynamics on social media, establishing a continuous monitoring mechanism that reduces feedback time to within one month. The “employee innovation capability” indicator can be broken down into quantifiable sub-indicators, such as the number of new ideas proposed and the proportion of innovation outcomes contributing to the company.

Second, a group-level data aggregation platform should be established to centrally integrate non-financial data from all brands and regions, enabling real-time sharing and analysis. This will accelerate the provision of evaluation feedback and allow timely guidance for operational improvements.

7.4 Expand the Application Scenarios of Performance Results and Strengthen Incentive Orientation

To address the limited application of performance results, Anta Sports should broaden the usage scenarios of Balanced Scorecard outcomes and enhance their incentive orientation by closely integrating them with the reward system [4].

First, a stronger linkage should be established between performance results and employee training, promotion, and career development. Capability gaps identified through performance evaluations should inform the design of personalized training programs, and performance outcomes should serve as a key factor in promotion decisions, thereby encouraging employees to continuously improve their competencies.

Second, the incentive intensity for long-term strategic indicators should be increased. Indicators such as “improvement in the proportion of overseas market revenue” and “effectiveness of multi-brand collaboration” should be incorporated into the core reward framework. Long-term incentive schemes (e.g., three-year tenure bonuses) should be introduced to motivate individual brands to focus on the group’s long-term strategic direction, thereby supporting the effective implementation of globalization and multi-brand synergy strategies [5].

8. Conclusions and Prospects

8.1 Research Conclusions

This study takes Anta Sports Products Group as a case to examine the application of the Balanced Scorecard in the context of multi-brand operations and globalization. It finds that the Balanced Scorecard, through its four perspectives—financial, customer, internal processes, and learning and growth—effectively translates the group’s overall strategy into specific, actionable performance indicators. This enables each brand to conduct performance management aligned with the group’s strategic objectives. As a result, Anta Sports has achieved notable improvements in financial performance, market competitive positioning, operational efficiency, and long-term development capabilities.

These findings demonstrate that the Balanced Scorecard serves not only as an effective tool for short-term performance management but also as a powerful mechanism to help enterprises achieve long-term strategic goals.

Nevertheless, the study also identifies several limitations in its application, including unreasonable allocation of indicator weights, insufficient differentiation in globalization-related evaluation, difficulties in quantifying non-financial indicators, and the overly narrow application of performance results. Enterprises can further enhance the effectiveness of the Balanced Scorecard by adjusting indicator weights, developing region-specific indicators, improving the quantification of non-financial metrics, and strengthening incentive mechanisms.

8.2 Research Prospects

With the continued globalization and digital transformation of the sporting goods industry, Anta Sports and other multi-brand enterprises should place greater emphasis on data integration and analysis when applying the Balanced Scorecard. This will improve both the speed and accuracy of performance evaluation.

Furthermore, the indicator system can be further refined according to the unique characteristics of the industry by incorporating factors such as event marketing and product technological innovation. Such customization will make the Balanced Scorecard more suitable for the specific context of the sporting goods sector.

Other enterprises may also draw lessons from Anta’s experience, adapting the Balanced Scorecard to their own circumstances and integrating it with other management accounting tools. By doing so, they can build

more flexible and effective performance evaluation systems that achieve a close alignment between corporate strategy and operational performance.

References

- [1] Kaplan, R. S., & Norton, D. P. (1992). The Balanced Scorecard—Measures that drive performance. *Harvard Business Review*, 70(1), 71–79.
- [2] Luo, Y., & Sun, Y. (2019). Integration of the Balanced Scorecard and internal control in advanced manufacturing enterprises. *Finance & Accounting Monthly*, (5), 44–51.
- [3] Sun, Y. (2013). A brief analysis of the application of the Balanced Scorecard in enterprises. *China Certified Public Accountant*, (7), 78–81.
- [4] Zhang, Y. (2009). Application of the Balanced Scorecard in enterprise performance management. *Contemporary Economy*, (2), 74–75.
- [5] Li, Y., & Liu, H. (2010). Analysis of the application of BSC in performance appraisal management of Chinese enterprises. *Enterprise Technology Development*, 29(1), 162–164. <https://doi.org/10.14165/j.cnki.hunansci.2010.01.067>

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Safety-First Pension Funds, Governance, and Market Risk: Evidence from China's A-Share Market

Yun'e Xie*

Hebei Oriental University, Langfang 065001, China

*Corresponding author: Yun'e Xie.

Abstract

This paper examines how Chinese public pension funds affect corporate governance and market risk in a state-capitalist setting. Using A-share firms from 2015 to 2024, we show that pension funds display an ESG screening effect *ex ante* and are followed by higher post-investment ESG scores. The effect, however, is concentrated in the governance pillar and is more than twice as large in SOEs as in private firms. Despite these governance gains, pension ownership is not associated with lower stock-return volatility. We interpret this pattern as governance-risk decoupling: pension funds appear to promote visible, compliance-oriented governance upgrades, but their small stakes and the retail-dominated market limit their ability to stabilize prices. The findings refine the universal-owner view by showing that, in transitional economies, state-backed patient capital may function more as administrative legitimacy than as a direct mechanism of market-risk reduction.

Keywords

public pension funds, ESG, symbolic management, governance-risk decoupling, state capitalism, China A-share market

1. Introduction

In sustainable finance, institutional investors are often viewed as stewards that internalize externalities and reduce portfolio risk through long-horizon monitoring [1, 2, 3, 4]. That logic is less straightforward in China's A-share market, where concentrated ownership, state influence, and weaker investor protection shape firm behavior [5, 6, 7]. Chinese public pension funds are especially informative because they combine market participation with a Safety-First mandate centered on asset safety and value preservation [8, 9].

We argue that these funds operate less as conventional market disciplinarians than as carriers of administrative legitimacy. Because their holdings are extremely small, their influence is unlikely to come from voting power alone. Instead, pension-fund entry may signal official scrutiny or policy relevance, encouraging managers—especially in SOEs—to strengthen visible governance practices and regulatory compliance rather than undertake deeper changes in risk structure [10, 11, 12, 13].

Using Chinese A-share firms from 2015 to 2024, we document governance-risk decoupling. Pension funds are more likely to hold firms with stronger prior ESG performance, and pension ownership is followed by higher ESG scores. Yet the post-investment effect is concentrated in governance: we find no comparable

improvement in E or S, and the G effect is more than twice as large in SOEs as in private firms. This pattern differs from the conventional ESG-risk literature, which predicts that better governance should lower volatility [14, 15, 16, 17, 18].

The average pension stake in our sample is only about 0.05%, too small to support a standard monitoring story based on concentrated ownership or activism [19, 20]. We therefore interpret the governance response as largely symbolic and compliance-oriented: firms can quickly improve disclosure routines, board procedures, and internal controls that are visible to regulators, even if those changes do not materially alter market risk.

Consistent with this interpretation, pension ownership does not significantly reduce future stock-return volatility. Chinese public pension funds thus appear to generate regulatory safety more readily than market safety.

This paper contributes in three ways. First, it refines the universal-owner perspective by showing that patient capital in a state-capitalist system may operate through administrative signaling rather than market discipline [1, 3, 4]. Second, it identifies state-sector heterogeneity: SOEs respond far more strongly than private firms, consistent with their greater sensitivity to politically salient investors [6, 7, 21]. Third, it shows that visible governance improvement and substantive market-risk reduction can diverge in transitional economies [10, 12].

2. Literature Review

2.1 Institutional Investors, Public Pension Funds, and Corporate Governance in China

Institutional investors are often seen as external monitors, but their effectiveness depends on incentives, horizons, ownership concentration, and legal protection [19, 22]. In China, monitoring takes place in an environment of concentrated ownership, controlling-shareholder power, and relatively weak minority-investor protection, so political connections and administrative hierarchy matter alongside equity stakes [5, 6, 7, 21, 23]. This is especially relevant for public pension funds, which combine a long horizon with a prudential public mandate and typically small holdings. Their presence may therefore operate as both an investment decision and a reputational or administrative signal, especially for SOEs. Recent evidence suggests that institutional shareholders can improve ESG outcomes in China, but the strength and substance of the effect remain context dependent [24].

2.2 ESG Performance, Governance Quality, and the Conventional Expectation of Risk Reduction

A broad literature links ESG performance to firm value, financing conditions, and risk. Nonfinancial disclosure can lower the cost of equity, and sustainability performance seems most relevant when it concerns financially material issues [25, 26, 27]. For governance in particular, the key mechanism is information opacity: when disclosure is poor and bad news is hoarded, firms become more vulnerable to volatility [28, 29]. Consistent with this view, stronger ESG or CSR performance is associated with lower downside risk in several settings, including China, and ESG engagement can also mitigate risk [14, 15, 16, 17, 18, 30, 31]. For public pension funds, this literature implies both ex ante screening for better ESG firms and post-investment pressure for further improvement.

2.3 Symbolic Management, State Capitalism, and Governance-Risk Decoupling

A more skeptical perspective comes from symbolic management and institutional decoupling. Organizations may adopt formal structures to gain legitimacy without substantially changing operating practices [10]. In governance settings, firms often implement visible reforms that satisfy outside audiences while leaving core incentives intact [11, 12]. This concern is especially salient in China, where firms respond not only to markets but also to regulators and politically salient investors. Governance indicators are comparatively easy to standardize, disclose, and verify, so they are well suited to legitimacy-oriented adjustment. Related work on CSR reporting in China similarly shows that sustainability disclosure can serve as symbol or substance depending on institutional context [32].

2.4 Research Gap and Hypotheses Development

These literatures generate competing expectations. One predicts that long-horizon institutional investors and stronger ESG performance should improve governance and reduce downside risk; the other suggests that visible governance reform in a state-dominated market may remain partly symbolic [10, 12, 15, 25, 31, 32]. Existing work has rarely integrated these perspectives for Chinese public pension funds, which combine a prudential mandate with operation in a politically structured market [6, 7, 8]. We therefore ask whether pension funds screen on ESG, whether their influence is concentrated in G rather than E and S, and whether any governance gains translate into lower stock-return volatility.

3. Institutional Background and Hypotheses Development

3.1 The “Safety-First” Mandate and Risk Screening

Chinese public pension funds invest under a Safety-First mandate that stresses market-based operation, diversification, asset safety, and value preservation [8]. In an information-friction environment, this mandate gives managers reason to prefer firms with stronger ESG profiles, which prior research associates with lower risk and a better information environment [14, 15, 16, 17, 18, 30, 33]. We therefore expect an ex ante screening effect.

H1: Pension funds are more likely to hold firms with higher initial ESG performance (the screening effect).

3.2 Administrative Legitimacy and Governance Signaling

Traditional investor influence is usually framed in terms of voting, engagement, or exit, but these channels are constrained here by tiny ownership stakes. In China’s state-capitalist system, pension-fund entry may instead serve as a signal of administrative legitimacy. Because SOE managers are embedded in political evaluation and personnel systems, they are especially likely to respond to such signals by strengthening visible governance arrangements—disclosure rules, board procedures, and internal compliance—rather than by making slower, costlier changes in E or S. Recent Chinese evidence that institutional investors can affect ESG outcomes through site visits and other information channels is consistent with this view [6, 7, 21, 34].

H2: Pension fund ownership is positively associated with the Governance (G) scores of portfolio firms, and this effect is significantly stronger in state-owned enterprises (SOEs).

3.3 The Governance-Risk Decoupling

The crucial question is whether these governance gains amount to real market stabilization. Under the standard opacity channel, better governance should reduce bad-news hoarding and therefore lower volatility [15, 28, 29]. But symbolic-management theory predicts that firms may adopt monitorable, politically legible reforms without changing underlying risk drivers [10, 11, 12, 32]. Given the tiny pension holdings and the importance of noise trading in China’s market, we expect visible governance improvement without a statistically meaningful decline in subsequent volatility [35].

H3: Pension fund ownership is positively associated with formal governance improvement, but not with a statistically significant reduction in subsequent stock-return volatility (the governance-risk decoupling hypothesis).

4. Research Design

4.1 Sample Selection and Data Source

We study Chinese A-share listed firms from 2015 to 2024. The sample begins in 2015 because the State Council issued the Measures for the Administration of Investment of Basic Pension Insurance Funds in August 2015, which established the regulatory framework for the entrusted investment of provincial basic pension insurance fund balances. Actual entrusted operation began in December 2016, so the early years of the sample capture the institutional transition surrounding policy rollout rather than the first appearance of public pension capital in equities. This distinction matters because the National Social Security Fund had already been

permitted to invest in listed stocks and securities investment funds under the 2001 interim rules [8]. Pension holdings are manually collected from the “Top Ten Shareholders” and “Top Ten Tradable Shareholders” sections of annual and semiannual reports in CSMAR. Our measure captures publicly disclosed holdings attributable to the National Social Security Fund and, where separately identifiable, basic pension insurance funds. We do not include enterprise annuity accounts or other supplementary pension arrangements, because the theoretical mechanism in this paper concerns state-backed public pension capital and its visibility-based legitimacy effects.

ESG data come from the Huazheng (Sino-Securities) ESG database, which is widely used in Chinese ESG research and, crucially, provides both a composite score and separate E, S, and G pillars [16, 18, 36]. We exclude ST and *ST firms, exclude financial firms, drop observations with missing key variables, and winsorize continuous variables at the 1st and 99th percentiles. The final sample contains 19,179 firm-year observations.

4.2 Variable Definitions

4.2.1 Dependent Variables

Our outcomes fall into three groups. For ESG responses, we use the Huazheng composite ESG score and the E, S, and G pillar scores. The pillar decomposition is central to the paper because the theory predicts a concentrated governance response rather than symmetric improvement across ESG dimensions.

For market consequences, we examine future annualized stock-return volatility and, in robustness tests, crash-risk measures, following the bad-news-hoarding literature [15, 18, 28, 29]. We also report ROE as a supplementary accounting outcome.

4.2.2 Independent Variable

The key explanatory variable is *Pension_Ratio*, defined as the year-end shareholding ratio of publicly disclosed public pension funds in listed firms’ shareholder lists, specifically holdings attributable to the National Social Security Fund and, where separately identifiable, basic pension insurance funds. Because the mechanism in this paper is visibility based, disclosed rather than hidden exposure is the theoretically relevant measure. Enterprise annuity accounts and other supplementary pension arrangements are excluded from this measure.

4.2.3 Control Variables

Control variables follow the ESG, crash-risk, and Chinese corporate-governance literature and include firm size, leverage, ROE, Tobin’s Q, SOE status, and analyst coverage [7, 15, 16, 29]. These variables capture differences in fundamentals, information environment, and political sensitivity. Unless otherwise stated, regressions include fixed effects and firm-clustered standard errors.

4.3 Empirical Models

4.3.1 Testing the Screening Effect (H1)

To test H1, we regress *Pension_Ratio* on lagged *ESG_Score* with industry and year fixed effects. A positive coefficient on *L.ESG_Score* indicates that pension funds screen toward firms with stronger prior ESG performance.

$$Pension_{i,t} = \alpha_0 + \beta_1 ESG\_score_{i,t-1} + \gamma Controls_{i,t-1} + \lambda_j + \delta_t + \epsilon_{i,t} \quad (1)$$

4.3.2 Testing the Governance Watchdog Effect (H2)

To test H2, we regress future ESG outcomes on lagged pension ownership, alternating the dependent variable across the composite ESG score and the three pillar scores. A positive coefficient for G, especially without comparable effects for E or S, is consistent with administrative-legitimacy-based governance upgrading. Firm and year fixed effects are included.

$$Score_{i,t} = \alpha_0 + \beta_1 Pension_{i,t-1} + \gamma Controls_{i,t-1} + \mu_i + \delta_t + \epsilon_{i,t} \quad (2)$$

4.3.3 Testing Economic Consequences (H3)

To test H3, we regress future volatility on lagged pension ownership. Under the governance-risk decoupling hypothesis, the key prediction is that pension ownership improves formal governance but does not have a statistically significant negative effect on subsequent volatility. ROE is reported only as a supplementary accounting outcome. Under the governance-risk decoupling hypothesis, the key prediction is that pension ownership improves formal governance but does not have a statistically significant negative effect on subsequent volatility.

$$Outcome_{i,t+1} = \alpha_0 + \beta_1 Pension_{i,t-1} + \gamma Controls_{i,t} + \mu_i + \delta_t + \epsilon_{i,t} \quad (3)$$

4.4 Endogeneity Strategy: Propensity Score Matching (PSM)

Selection is an important concern because pension funds invest in only a small subset of A-share firms. We therefore implement propensity score matching following Rosenbaum and Rubin [37].

Using firm size, leverage, Tobin's Q, and SOE status, we estimate the probability of pension investment and match treated firms to untreated firms through 1:3 nearest-neighbor matching. We then re-estimate the post-investment ESG regressions on the matched sample to assess whether the governance-related results survive observable balancing.

5. Empirical Results Analysis

5.1 Descriptive Statistics

Table 1 summarizes the sample. *Pension_Ratio* has a mean of 0.0536% and a median of zero, showing that pension participation is highly concentrated and that typical holdings are too small to imply strong control rights. The descriptive statistics also show higher average G than E scores, which is consistent with the idea that governance practices are easier to standardize and document than deeper operational change [19, 20].

Table 1: Descriptive Statistics

Variable	N	Mean	SD	Min	Median	Max
Pension_Ratio	19,179	0.0536	0.2656	0.0000	0.0000	1.8138
ESG_Score	19,179	74.31	5.33	56.51	74.27	86.77
G_Score	19,179	79.54	5.95	19.60	80.47	97.33
E_Score	19,179	62.83	7.96	31.08	62.02	92.30
S_Score	19,179	77.62	8.26	9.77	78.17	100.00
Size	19,179	22.74	1.31	20.07	22.54	26.39
Leverage	19,179	0.4184	0.1910	0.0558	0.4145	0.8938
ROE	19,179	0.0799	0.1123	-0.6393	0.0831	0.3614
TobinQ	19,179	2.3584	1.6686	0.8130	1.8370	10.8681
Volatility	19,179	0.4422	0.1279	0.1928	0.4350	2.1252
Analyst_Log	19,179	2.0042	0.9052	1.0000	2.0000	4.0000
SOE	19,179	0.3001	0.4583	0.0000	0.0000	1.0000

Notes: This table presents descriptive statistics for the final sample of Chinese A-share listed firms from 2015 to 2024. The final sample contains 19,179 firm-year observations. Detailed variable definitions are reported in Appendix A.

5.2 Pension Fund ESG Allocation Preference (Testing H1)

Table 2 supports H1. The coefficient on *L.ESG_Score* is positive and significant ($\beta = 0.0019$, $t = 3.41$), indicating that pension funds tilt toward firms with stronger prior ESG performance. This pattern is consistent with prudent screening under a Safety-First mandate. The SOE coefficient is also positive and significant, suggesting state-sector affinity in portfolio allocation, although that result may reflect several channels, including lower political uncertainty and greater institutional compatibility [5, 6, 7].

5.3 Baseline Post-Investment ESG Response (Initial Test of H2)

Table 3 shows that lagged pension ownership is positively associated with subsequent ESG scores ($\beta = 0.4761$, $t = 2.16$). This result suggests a post-investment response, but it does not by itself identify the relevant ESG dimension. Given the tiny average holding, the estimate is more consistent with visible but limited influence than with strong ownership-based intervention [19, 20, 24, 34].

Table 2: Pension Fund Preference (Testing H1)

Variable	(1) Pension Ratio
L.ESG Score	0.0019***
	(3.41)
Size	-0.0066**
	(-2.15)
Leverage	-0.0104
	(-0.57)
TobinQ	0.0023
	(1.34)
SOE	0.0390***
	(4.43)
Year FE	Yes
Observations	19,179
R-squared	0.0127

Notes: The dependent variable is *Pension_Ratio*. The main independent variable is *L.ESG_Score* (lagged one year). The regression includes Industry and Year fixed effects. T-statistics based on robust standard errors clustered at the firm level are reported in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 3: Baseline Post-Investment ESG Response

Variable	(1) ESG Score
L.Pension Ratio	0.4761**
	(2.16)
Size	1.9029***
	(8.68)
Leverage	-6.7594***
	(-9.51)
TobinQ	-0.1524***
	(-3.30)
SOE	0.0828
	(0.18)
Firm FE	Yes
Year FE	Yes
Observations	19,179
R-squared (Within)	0.0374

Notes: The dependent variable is the ESG composite score (*ESG_Score*). The key independent variable is the one-year lagged pension fund ownership ratio (*L.Pension_Ratio*). All models include Firm and Year fixed effects. T-statistics based on robust standard errors clustered at the firm level are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

At the same time, the interpretation of this result requires care. The average pension stake in the sample is extremely small, which makes a conventional control-rights explanation less persuasive. It is therefore difficult to attribute the ESG response to strong ownership-based intervention of the type usually associated with concentrated institutional monitoring [19, 20]. A more plausible reading is that even small but publicly visible pension positions can induce firms to respond in ways that improve observable ESG metrics. This interpretation is consistent with recent Chinese evidence showing that institutional investors can affect firms' ESG-related outcomes, although the precise transmission channels may vary across settings [24, 34].

Importantly, Table 3 should be read as a baseline reduced-form result. It establishes that pension ownership is associated with higher subsequent ESG performance, but it does not by itself identify which ESG pillar drives that effect. This distinction matters for the argument of this paper. Our theory predicts that the post-investment response should be concentrated in governance rather than evenly distributed across E, S, and G, and that the governance response should be stronger in SOEs. Those sharper tests are conducted in the next chapter through pillar-level decomposition and ownership heterogeneity analysis.

5.4 Economic Consequences: Profitability and Risk Exposure (Testing H3)

The core test of Hypothesis 3 concerns whether pension-fund-related governance improvement translates into lower subsequent market risk. In line with the research design set out earlier, the main outcome for H3 is future stock return volatility, while ROE is treated as a supplementary accounting outcome.

The central H3 test appears in Table 5. The coefficient on *L.Pension_Ratio* is negative but statistically insignificant ($\beta = -0.0047$, $t = -1.46$), so pension ownership does not measurably reduce future stock-return volatility. This null result contrasts with the conventional ESG-risk view, under which better governance should improve transparency and lower downside risk [15, 16, 17, 18, 28, 29].

Table 4 provides supplementary evidence on profitability. *L.Pension_Ratio* is also insignificant in the ROE regression ($\beta = -0.0028$, $t = -0.90$), suggesting no detectable short-run operating effect. Taken together, the evidence supports governance-risk decoupling: pension ownership is associated with visible governance-oriented adjustment, but not with statistically significant improvements in market stabilization or near-term profitability [10, 12, 32].

Table 4: Impact on Corporate Profitability

Variable	(1) ROE
<i>L.Pension_Ratio</i>	-0.0028
	(-0.90)
Size	0.0778***
	(14.33)
Leverage	-0.2941***
	(-13.45)
TobinQ	0.0218***
	(21.77)
SOE	-0.0175
	(-1.52)
Firm FE	Yes
Year FE	Yes
Observations	19,179
R-squared (Within)	0.0747

Notes: The dependent variable is Return on Equity (ROE). The key independent variable is the one-year lagged pension fund ownership ratio (*L.Pension_Ratio*). All models include Firm and Year fixed effects. T-statistics based on robust standard errors clustered at the firm level are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Table 5: Impact on Stock Price Volatility

Variable	(1) Volatility
<i>L.Pension_Ratio</i>	-0.0047
	(-1.46)
Size	-0.0253***
	(-6.47)
Leverage	0.1084***
	(8.07)
TobinQ	0.0246***
	(18.48)
SOE	0.0061
	(0.74)
Firm FE	Yes
Year FE	Yes
Observations	19,179
R-squared (Within)	0.0816

Notes: The dependent variable is annualized stock return volatility (Volatility). The key independent variable is the one-year lagged pension fund ownership ratio (*L.Pension_Ratio*). All models include Firm and Year fixed effects. T-statistics based on robust standard errors clustered at the firm level are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

6. Mechanism and Heterogeneity Analysis

The baseline results imply a decoupled pattern: pension ownership raises observed ESG scores without reducing volatility. We therefore examine whether the effect is concentrated in governance, whether it varies by ownership type, whether it operates through analyst attention, and whether it survives matching.

6.1 Pillar Decomposition and State-Sector Heterogeneity

Table 6 shows that the post-investment ESG response is concentrated in governance. *L.Pension_Ratio* is insignificant for E and S, but positive and highly significant for G ($\beta = 0.9577$, $t = 4.59$). This asymmetry fits the argument that governance practices are more visible, standardized, and easier to adjust quickly than environmental or social outcomes, making them especially suitable for legitimacy-oriented responses [10, 12, 32].

The ownership split sharpens the result. Pension ownership is positively related to G in both private firms and SOEs, but the coefficient is much larger in SOEs (1.2889 versus 0.4853). Because holdings are small in both subsamples, this gap is difficult to explain by financial monitoring alone and is more consistent with stronger responsiveness to state-backed signals inside the state sector [6, 7, 21]. These results support H2.

Table 6: Mechanism and Heterogeneity – Decomposing ESG and Ownership Structure

Variable	E Score	S Score	G Score	(4) G Score	(5) G Score
Sample	Full	Full	Full	Private (Non-SOE)	State-Owned (SOE)
L.Pension_Ratio	-0.1347 (-0.40)	0.1033 (0.31)	0.9577*** (4.59)	0.4853** (2.18)	1.2889*** (3.60)
Size	1.9490*** (5.64)	1.8615*** (5.31)	1.7927*** (7.37)	1.4828*** (5.38)	1.5029*** (3.18)
Leverage	-4.3359*** (-4.12)	-1.2495 (-1.11)	-10.105*** (-12.44)	-10.242*** (-11.18)	-9.6737*** (-6.01)
TobinQ	-0.2454*** (-3.58)	0.0705 (1.02)	0.0563 (1.11)	0.1143** (1.99)	-0.13 (-1.16)
SOE	-0.0494 (-0.07)	0.2978 (0.72)	0.3486 (0.62)	- -	- -
Observations	19,179	19,179	19,179	13,424	5,755
R-squared (Within)	0.068	0.0347	0.0258	0.0332	0.0094
Firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Notes: This table reports post-investment ESG regressions by pillar and ownership type. Columns (1)–(3) use E_Score, S_Score, and G_Score as the dependent variables, respectively. Columns (4) and (5) report G_Score regressions for private firms and SOEs, respectively. The regression includes Firm and Year fixed effects. T-statistics based on robust standard errors clustered at the firm level are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

6.2 Limited Evidence of a Spotlight Effect

A potential alternative explanation is a spotlight effect: pension ownership may attract analyst attention and thereby improve behavior through greater outside scrutiny. Table 7 provides little support for this mechanism. Although the coefficient on *L.Pension_Ratio* is positive, it is statistically insignificant ($\beta = 0.0373$, $t = 1.37$), suggesting that pension entry does not systematically expand analyst coverage. This does not rule out all forms of external attention, but it weakens the view that the governance effect mainly reflects an analyst-attention channel.

Table 7: Limited Evidence of a Spotlight Effect

Variable	(1) Analyst_Log
L.Pension_Ratio	0.0373 (1.37)
Controls	Yes
Firm FE	Yes
Year FE	Yes
Observations	19,179
R-squared (Within)	0.0002

Notes: The dependent variable is the natural logarithm of analyst coverage (*Analyst_Log*). The key independent variable is the one-year lagged pension fund ownership ratio (*L.Pension_Ratio*). All models include Firm and Year fixed effects. T-statistics based on robust standard errors clustered at the firm level are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

6.3 Robustness Check: Propensity Score Matching (PSM)

We next address observable selection with propensity score matching. Treated firms are matched to untreated firms using 1:3 nearest-neighbor matching on firm size, leverage, Tobin's Q, and SOE status [37].

Table 8 shows that the positive post-investment ESG association remains in the matched sample ($\beta = 0.5124$, $t = 2.31$; 11,204 observations). Because the matched regression uses the composite ESG score, it confirms the baseline post-investment effect rather than the G-specific mechanism by itself. Read together with Table 6, however, the evidence still supports a governance-centered and state-contingent response that is robust to observable selection.

Table 8: Robustness Check – Propensity Score Matching (PSM)

Variable	(1) ESG Score
L.Pension Ratio	0.5124**
	(2.31)
Controls	Yes
Firm FE	Yes
Year FE	Yes
Observations (Matched)	11,204
R-squared (Within)	0.0392

*Notes: This table presents the robustness check using a Propensity Score Matching (PSM) sample. We employ a 1:3 nearest-neighbor matching based on firm size, leverage, Tobin's Q, and SOE status to construct a control group. The dependent variable is the ESG composite score (ESG_Score). All models include Firm and Year fixed effects. T-statistics based on robust standard errors clustered at the firm level are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.*

7. Conclusion and Implications

This paper studies Chinese public pension funds as state-backed long-horizon investors in a state-capitalist market. Using A-share firms from 2015 to 2024, we find three main results: pension funds screen toward firms with stronger prior ESG performance; pension ownership is followed by higher subsequent ESG scores, but the effect is concentrated in governance and is much stronger in SOEs; and these governance gains are not accompanied by statistically significant declines in volatility or short-run ROE. The overall pattern is consistent with governance-risk decoupling.

These findings qualify the universal-owner view. In this setting, the average disclosed pension stake is too small to support a strong control-rights interpretation [19, 20]. The evidence instead suggests that pension ownership works partly through administrative legitimacy and public visibility. For SOEs in particular, pension entry may signal political salience and induce more formal, compliance-oriented governance responses than deep changes in the underlying drivers of risk [7, 10, 12].

The study contributes by linking institutional investment, Chinese corporate governance, and symbolic management. It shows that state-backed patient capital does not necessarily operate like market-driven institutions, that ownership type conditions firm responses, and that better formal governance scores need not imply lower market risk.

The policy implication is that regulators and pension managers should not treat high governance scores as sufficient evidence of effective risk control. Evaluation frameworks may need to complement process-based metrics with outcome-based measures such as realized volatility, downside exposure, and drawdown resilience. For firms, especially SOEs, visible governance compliance may attract state-backed long-horizon capital, but it is unlikely to reduce market risk without deeper improvements in disclosure quality, internal controls, and operating practice.

Several limitations remain. The design is observational despite the use of lagged regressions and PSM; our ownership measure captures publicly disclosed rather than total exposure; ESG ratings may register formal compliance more readily than deep governance quality; and the paper cannot directly observe informal engagement or administrative coordination. Future research can exploit quasi-natural experiments, examine tail-risk outcomes in more detail, and compare pension funds with insurers, mutual funds, and foreign institutions.

References

- [1] Dyck, A., Lins, K. V., Roth, L., & Wagner, H. F. (2019). Do institutional investors drive corporate social responsibility? International evidence. *Journal of Financial Economics*, 131(3), 693–714.
- [2] Krueger, P., Sautner, Z., & Starks, L. T. (2020). The importance of climate risks for institutional investors. *Review of Financial Studies*, 33(3), 1067–1111.
- [3] Pedersen, L. H., Fitzgibbons, S., & Pomorski, L. (2021). Responsible investing: The ESG-efficient frontier. *Journal of Financial Economics*, 142(2), 572–597.
- [4] Matos, P. (2020). ESG and responsible institutional investing around the world: A critical review. CFA Institute Research Foundation.
- [5] Allen, F., Qian, J., & Qian, M. (2005). Law, finance, and economic growth in China. *Journal of Financial Economics*, 77(1), 57–116.
- [6] Fan, J. P. H., Wong, T. J., & Zhang, T. (2007). Politically connected CEOs, corporate governance, and post-IPO performance of China's newly partially privatized firms. *Journal of Financial Economics*, 84(2), 330–357.
- [7] Jiang, F., & Kim, K. A. (2020). Corporate governance in China: A survey. *Review of Finance*, 24(4), 733–772.
- [8] State Council of the People's Republic of China. (2015, August 23). Measures for the administration of investment in basic pension insurance funds [in Chinese].
- [9] Hammond, P. B., Maurer, R., & Mitchell, O. S. (Eds.). (2023). *Pension funds and sustainable investment: Challenges and opportunities*. Oxford University Press.
- [10] Meyer, J. W., & Rowan, B. (1977). Institutionalized organizations: Formal structure as myth and ceremony. *American Journal of Sociology*, 83(2), 340–363.
- [11] Westphal, J. D., & Zajac, E. J. (1994). Substance and symbolism in CEOs' long-term incentive plans. *Administrative Science Quarterly*, 39(3), 367–390.
- [12] Westphal, J. D., & Zajac, E. J. (1998). The symbolic management of stockholders: Corporate governance reforms and shareholder reactions. *Administrative Science Quarterly*, 43(1), 127–153.
- [13] Westphal, J., & Park, S. H. (2020). *Symbolic management: Governance, strategy, and institutions*. Oxford University Press.
- [14] Albuquerque, R., Koskinen, Y., & Zhang, C. (2019). Corporate social responsibility and firm risk: Theory and empirical evidence. *Management Science*, 65(10), 4451–4469.
- [15] Kim, Y., Li, H., & Li, S. (2014). Corporate social responsibility and stock price crash risk. *Journal of Banking & Finance*, 43, 1–13.
- [16] Feng, J., Goodell, J. W., & Shen, D. (2022). ESG rating and stock price crash risk: Evidence from China. *Finance Research Letters*, 46, Article 102476.
- [17] Liu, D., Gu, K., & Hu, W. (2023). ESG performance and stock idiosyncratic volatility. *Finance Research Letters*, 58, Article 104393.
- [18] Luo, W., Tian, Z., Fang, X., & Deng, M. (2024). Can good ESG performance reduce stock price crash risk? Evidence from Chinese listed companies. *Corporate Social Responsibility and Environmental Management*, 31(3), 1469–1492.
- [19] Chen, X., Harford, J., & Li, K. (2007). Monitoring: Which institutions matter? *Journal of Financial Economics*, 86(2), 279–305.
- [20] Becht, M., Franks, J., Mayer, C., & Rossi, S. (2009). Returns to shareholder activism: Evidence from a clinical study of the Hermes UK Focus Fund. *Review of Financial Studies*, 22(8), 3093–3129.
- [21] Li, H., & Zhou, L.-A. (2005). Political turnover and economic performance: The incentive role of personnel control in China. *Journal of Public Economics*, 89(9–10), 1743–1762.

- [22] Ferreira, M. A., & Matos, P. (2008). The colors of investors' money: The role of institutional investors around the world. *Journal of Financial Economics*, 88(3), 499–533.
- [23] Jiang, G., Lee, C. M. C., & Yue, H. (2010). Tunneling through intercorporate loans: The China experience. *Journal of Financial Economics*, 98(1), 1–20.
- [24] Sun, Y., & Zhao, Z. (2024). Responsible investment: Institutional shareholders and ESG performance. *Pacific-Basin Finance Journal*, 85, Article 102357.
- [25] Dhaliwal, D. S., Li, O. Z., Tsang, A., & Yang, Y. G. (2011). Voluntary nonfinancial disclosure and the cost of equity capital: The initiation of corporate social responsibility reporting. *Accounting Review*, 86(1), 59–100.
- [26] El Ghouli, S., Guedhami, O., Kwok, C. C. Y., & Mishra, D. R. (2011). Does corporate social responsibility affect the cost of capital? *Journal of Banking & Finance*, 35(9), 2388–2406.
- [27] Khan, M., Serafeim, G., & Yoon, A. (2016). Corporate sustainability: First evidence on materiality. *Accounting Review*, 91(6), 1697–1724.
- [28] Jin, L., & Myers, S. C. (2006). R² around the world: New theory and new tests. *Journal of Financial Economics*, 79(2), 257–292.
- [29] Hutton, A. P., Marcus, A. J., & Tehranian, H. (2009). Opaque financial reports, R², and crash risk. *Journal of Financial Economics*, 94(1), 67–86.
- [30] Jo, H., & Na, H. (2012). Does CSR reduce firm risk? Evidence from controversial industry sectors. *Journal of Business Ethics*, 110(4), 441–456.
- [31] Hoepner, A. G. F., Oikonomou, I., Sautner, Z., Starks, L. T., & Zhou, X. Y. (2024). ESG shareholder engagement and downside risk. *Review of Finance*, 28(2), 483–510.
- [32] Marquis, C., & Qian, C. (2014). Corporate social responsibility reporting in China: Symbol or substance? *Organization Science*, 25(1), 127–148.
- [33] Xu, Z., Liu, D., Li, Y., & Guo, F. (2025). ESG and stock price volatility risk: Evidence from Chinese A-share market. *North American Journal of Economics and Finance*, 75, Article 102277.
- [34] Jiang, Y., Wang, C., Li, S., & Wan, J. (2022). Do institutional investors' corporate site visits improve ESG performance? Evidence from China. *Pacific-Basin Finance Journal*, 76, Article 101884.
- [35] Foucault, T., Sraer, D. J., & Thesmar, D. J. (2011). Individual investors and volatility. *Journal of Finance*, 66(4), 1369–1406.
- [36] Wu, F., Zhu, B., & Tao, S. (2024). Can good ESG performance of listed companies reduce abnormal stock price volatility? Mediation effects based on investor attention. *PLOS ONE*, 19(9), Article e0307535.
- [37] Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

A Three-Dimensional “Government–Power–Carbon” Collaborative Optimization Framework for Energy-Intensive Industrial Parks

Mengyao Ren^{†,*} and Miao Shi[†]

Business School, Xi'an International Studies University, Xi'an 710128, China

[†]*These authors contributed equally to this work and share first authorship.*

^{*}*Corresponding author: Mengyao Ren.*

Abstract

Under the guidance of the “dual carbon” strategic goals, high-energy-consuming industrial parks, as the main carriers of industrial energy consumption and carbon emissions, are confronted with multiple challenges such as low energy utilization efficiency, increasingly strict carbon emission constraints, and slow policy responses. Traditional single governance approaches are unable to effectively coordinate the complex relationships among government regulation, power resource allocation, and carbon market mechanisms, thus there is an urgent need to establish a multi-dimensional mechanism coupling system optimization model. This paper focuses on the three-dimensional synergy of “government-electricity-carbon”, and constructs a collaborative optimization mechanism centered on the management organization layer, power system layer, and carbon flow network layer, exploring the path to achieve the maximization of resource allocation efficiency and the coordinated attainment of carbon emission control targets in high-energy-consuming industrial parks. To verify the feasibility of the mechanism, this paper takes the high-energy-consuming industrial park of Henan Luoyang Petrochemical as a case, and conducts simulation and comparative analysis based on on-site energy consumption and emission data. The research results show that this mechanism strengthens the dynamic regulatory role of government policies, enhances the response flexibility of the power system and the price discovery function of the carbon market. It not only provides theoretical support and practical paths for the low-carbon development of high-energy-consuming industrial parks, but also has significant practical significance and promotion value for achieving the carbon peak and carbon neutrality goals within the regional scope.

Keywords

high-energy-consuming industrial parks, “government-electricity-carbon” collaborative mechanism, low-carbon transformation, dual carbon goals

1. Introduction

With the in-depth promotion of China's strategy of “striving to reach the peak of carbon dioxide emissions by 2030 and striving to achieve carbon neutrality by 2060” (hereinafter referred to as “dual carbon”), the industrial field has become a key battlefield to achieve the goal of carbon emission reduction. As the core

carrier of industrial energy consumption and carbon emissions, the carbon emission intensity of high-energy-consuming parks is high, which has become the focus and difficulty in achieving the “dual carbon” goal. Take Henan Province, a large industrial province, as an example. According to the latest data, the total industrial carbon emissions account for more than 60% of the province [1], among which the proportion of carbon emissions in high-energy-consuming parks is as high as 45%, and the carbon intensity is significantly higher than the national average [2].

At present, these parks are generally caught in the development paradox of “increasing production means increasing carbon, and reducing carbon is afraid of losing money”: on the one hand, with the expansion of production scale, the rigid increase in energy consumption directly leads to an increase in carbon emissions [3]; on the other hand, the implementation of compulsory carbon emission reduction measures will significantly increase the operating costs of enterprises and affect economic benefits. This contradiction has become a profound problem that restricts the green and low-carbon transformation of China's industry, and an innovative solution is urgently needed.

In-depth analysis shows that the root cause of this dilemma is not only the relative lag of technical equipment, but more importantly, the systemic rupture in the governance system of high-energy-consuming parks. Specifically, it is manifested in the coordination failure of the three dimensions of “government” (management organization layer, responsible for policy formulation and transmission), “electricity” (power system layer, responsible for energy scheduling and implementation), and “carbon” (carbon flow network layer, responsible for carbon emission accounting and constraints), forming a triple rigid constraints: first, there is a significant attenuation of the policy transmission mechanism, and it is difficult to accurately implement policies such as emission reduction targets, carbon quota allocation and green electricity incentives formulated by the government [4]; second, the scheduling optimization of the power system mainly focuses on electricity prices and load signals, and generally ignores the key impact of carbon emission constraints [5]; third, carbon emission management still adopts static annual accounting method, lacks real-time monitoring and dynamic feedback capabilities, and cannot support accurate energy decision-making and policy adjustment. This state of independent governance and lack of linkage makes it difficult to find a systematic solution at the park level that takes into account the goals of economic development and carbon emission reduction.

In response to the above problems, this paper innovatively builds a three-dimensional collaborative closed-loop optimization mechanism of “management organization-power system-carbon flow network” (referred to as M-E-C), and establishes a refined energy and carbon emission response system through the organic integration of the three key links of “policy transmission - energy scheduling - carbon emission control”. This study selects Henan Luoyang Petrochemical High Energy Consumption Park as a typical empirical case, and carries out systematic simulation and comparative analysis based on multi-source data (policy text, energy consumption data, carbon emission data). At the technical level, integrating LSTM and random forest algorithms to achieve high-precision load prediction, and improve the NSGA-II algorithm to support multi-target scheduling optimization with carbon quota constraints; at the institutional level, design carbon intensity dynamic assessment and green power premium mechanism to promote the transformation of enterprises from “passive fulfillment” to “active emission reduction”.

Through this innovative mechanism, this research aims to realize the coordinated optimization of power system dynamic scheduling and carbon emission costs, which not only directly responds to the national “dual carbon” strategic needs, but also provides a green transformation path that is economically feasible, technically reliable and policy-effective for more energy-consuming parks. The research results have important theoretical value and practical guidance to promote low-carbon development in the industrial field.

2. Literature Review and Practical Analysis

2.1 Literature Review

2.1.1 The Evolution and Limitations from “Electricity-Carbon Synergy” to “Government-Electricity-Carbon” Three-Dimensional Governance

Driven by the “dual carbon” goal, high-energy-consuming parks have become a key node to realize the green and low-carbon transformation of industry due to their large energy consumption, strong carbon

emissions and concentrated load. The current research has gradually formed a research framework with “technical coordination” and “market coupling” as the main line, and has made remarkable progress in the bilateral coupling of power system optimization and carbon market mechanism.

In terms of the electricity-carbon synergy mechanism, Yi Jun et al. [6] proposed a synergistic control system that includes indirect emission reduction assessment and electricity carbon accounting, expanding the analytical boundaries of energy systems under carbon constraints. Xiao Bai et al. [7] incorporated carbon trading factors into the power dispatch model, preliminarily achieving dual-objective optimization of economy and emission reduction. Ji Xingquan et al. [8] emphasized the synergistic configuration of low-carbon and economic objectives. These studies theoretically advanced the establishment of the electricity-carbon coupling analysis framework, but most of their models were based on a single system optimization paradigm, did not involve the policy transmission mechanism among the governance entities of the park, and failed to fully consider the regulatory role of the institutional structure on the evolution path of carbon flows. When facing the “triple disconnection” situation of policy transmission absence, power dispatch carbon constraints failure, and carbon management response delay in high-energy-consuming parks, they still lacked structural solutions.

2.1.2 Research At the Park Level: Deepening and Bottlenecks of Market Mechanisms and Enterprise Behaviors

In the research at the park level, the electricity-carbon market mechanism and enterprises' emission reduction behaviors have become the focus. The existing studies present two main characteristics: Firstly, the design of market mechanisms has become the mainstream direction. Zhang Guoliang et al. [9] constructed a coupled model of electricity-carbon-green certificate market, verifying its role in improving resource allocation efficiency; Dong Jiuzhou et al. [10] pointed out that electricity-carbon trading needs to be coordinated with the green certificate mechanism to motivate enterprises to decarbonize; Li Xianshan et al. [11] proposed a quota and CCER offset mechanism, optimizing the carbon cost control path. However, most of the aforementioned studies focused on the optimization of the market level itself, neglecting the institutional constraints and incentive effects of management organizations (such as the government, park management committees) on enterprises' behaviors, and failing to deeply integrate the market mechanism with the policy system. Secondly, research on the response of enterprise behaviors has begun to attract attention. Zhang Lingrong et al. and Liu Peide et al. [12-17] used evolutionary game theory and closed-loop supply chain models to precisely depict the micro-decision-making mechanism of enterprises under carbon price signals. However, the conclusions of these models are often difficult to extrapolate to the macro system design at the park level, and their deep real-time linkage with the power system is still insufficient.

2.1.3 Limitations of the Two Main Lines of “Technological Synergy” and “Market Coupling”

It is worth noting that the existing research has formed two relatively clear main lines of “technical coordination” and “market coupling”, but both of them have the problem of missing dimensions:

At the level of technical coordination, Liu Hao et al. [18] put forward a coordinated scheduling model for landscape carbon storage to improve user satisfaction by dynamically constraining carbon emissions; Hong Mingzhi et al. [19] built a carbon reduction model of waste heat cascade utilization to verify the economic and environmental protection benefits of industrial multi-energy flow coordination. However, these studies do not sufficiently portray the dynamic coupling mechanism of power system and carbon flow, and do not consider the regulatory effect of policy transmission on the selection of technical paths.

In the direction of market coupling, Li Lingfang et al. [20] designed a micro-grid cluster electricity-carbon collaborative trading structure in the industrial park, which effectively reduced operating costs; Huang Ying et al. [21] built a multi-scale decision-making model of power generation group electricity-carbon. However, its optimization dimension is still limited to the main interaction at the enterprise or group level, which fails to solve the problem of incentive transmission fault of “policy-enterprise” in the park.

The deeper common limitation lies in the lack of cross-level coordination mechanisms. For example, although Xu Zhenhua et al. (2025)'s [22-23] source network load reserve coordination model adopts robust optimization to respond to new energy fluctuations, carbon quota constraints are still dominated by static annual accounting; Chen Huixin et al.'s [24] pollution reduction and carbon reduction multi-objective optimization proves the effectiveness of the NSGA-II algorithm, but does not quantify the impact of the policy

coordination of management (such as government-park-enterprises) on the technical path; Hou Jingyi et al. [25] put forward the concept of pollution reduction and carbon reduction coordination, but lacks the design of the “electricity-carbon intensity double constraint” dynamic conduction mechanism unique to high-energy-consuming parks.

2.1.4 Theoretical Innovation Points

In conclusion, although the current research has achieved significant results in the field of electrocarbon synergy, when dealing with the practical predicament of high-energy-consuming industrial parks where “increasing production leads to increased carbon emissions and reducing carbon emissions may result in losses”, there are still two core deficiencies: Firstly, the existing coupling mechanisms lack dynamic interactivity and policy transmission paths, resulting in a sluggish response from the government to enterprises regarding carbon reduction requirements, and the synergy effect between policies, scheduling, and carbon management has not been fully exerted; Secondly, the current research mainly focuses on theoretical exploration and model construction, lacking empirical studies and multi-source data support for high-energy-consuming industrial parks.

Based on this, this study aims to construct a three-dimensional dynamic optimization model that connects policy formulation, power implementation, and carbon emission control, consisting of a “management organization layer (M), power system layer (E), and carbon flow network layer (C).” The innovations of this paper are: in terms of theoretical framework, it introduces Social Network Analysis (SNA) to quantify “policy synergy,” bridging the research gap between “institutional and behavioral” aspects in traditional models from a governance perspective; in terms of technical methods, it integrates LSTM and random forest to achieve high-frequency load forecasting and improves the NSGA-II algorithm to support multi-objective optimization with dynamic carbon constraints, realizing real-time linkage between “source-grid-load-storage”; in terms of system integration, it strives to open up the key closed-loop path of “institutional-technology-market,” providing a theoretically closed-loop and practically implementable systematic solution to overcome the dilemma of three-dimensional fragmentation.

2.2 Practical Analysis

To gain a deeper understanding of the practical pathways, achievements, and bottlenecks in the coordinated governance of electricity and carbon emissions in high-energy-consuming industrial parks, this study selects representative industrial parks in Guangdong and Jiangsu provinces, which excel in policy mechanism innovation and technology integration, for in-depth case analysis. While these cases have achieved significant progress, they also reveal common challenges, providing important practical insights and optimization directions for constructing the three-dimensional coordinated governance framework in this paper.

2.2.1 Case One: A High-tech Industrial Development Zone in the Pearl River Delta of Guangdong Province

Currently, a deep coupling of the “electricity-carbon-green certificate” market has been achieved [26-28], deeply integrating park-level carbon quota trading, direct green electricity trading, and green certificate subscription mechanisms. Enterprises can sell excess carbon quotas through the park's internal trading platform, and the proceeds can be directly used to offset the premium for green electricity procurement or to purchase green certificates, forming a closed-loop incentive of “emission reduction-increased revenue-further emission reduction”. The park management committee, in collaboration with China Southern Power Grid and the carbon exchange, has established a unified 'Electricity-Carbon Coordinated Trading Platform,' achieving data interoperability and one-stop trading. At the same time, the park has introduced dynamic carbon price signals to guide load adjustments. When the ratio of real-time to expected carbon prices exceeds a set threshold, it will automatically trigger optimized operation strategies for high-energy-consuming adjustable loads (such as air conditioning systems), or increase energy storage discharge, reducing operational costs and carbon emissions during periods of high electricity and carbon prices. The proportion of enterprises participating in green electricity trading increased from 15% at the beginning of the pilot program to 45%, with green electricity consumption growing at an average annual rate of 32%. Some enterprises have increased their annual revenue by over one million yuan through carbon quota trading, effectively offsetting some of the transformation costs. The carbon emission intensity per unit of output in the industrial park has also decreased by about 12% within two years, and the green electricity penetration rate has increased to 25%. This model has also been promoted

by the National Development and Reform Commission as a typical experience due to its clear market-oriented incentive mechanism.

2.2.2 Case Two: A National-level Economic and Technological Development Zone in Southern Jiangsu Province

Nowadays, this development zone in Jiangsu Province has implemented a “management - technology - platform” tripartite model, led by the park management committee and established as the “Electricity and Carbon Smart Operation Center”. It integrates power grid dispatch data, enterprise energy consumption data, and online environmental monitoring data (CEMS) to build a unified “Electricity and Carbon Digitalization Platform”. This platform can already achieve multiple functions such as precise carbon profiling, coordinated optimization of source, grid, load, and storage, dynamic monitoring and early warning of carbon intensity, etc. By leveraging digital technology to enhance management coordination, the initiative provides the Management Committee, the power grid company, and key enterprises with a platform for data sharing and collaborative decision-making, thereby improving the efficiency of policy communication and implementation. The installed capacity of energy storage systems in the park has also doubled within two years, making them a vital source of flexibility.

2.2.3 Summary and Implications of Practical Analysis

Based on the practices of the leading parks in the two provinces, we can clearly see the progress and achievements of the two provinces: through market mechanism innovation (Guangdong) or technology platform integration (Jiangsu), significant results have been achieved in improving green electricity consumption, reducing carbon intensity, exploring the economics of energy storage, and improving management coordination, which has verified the feasibility and potential of electricity-carbon synergy optimization.

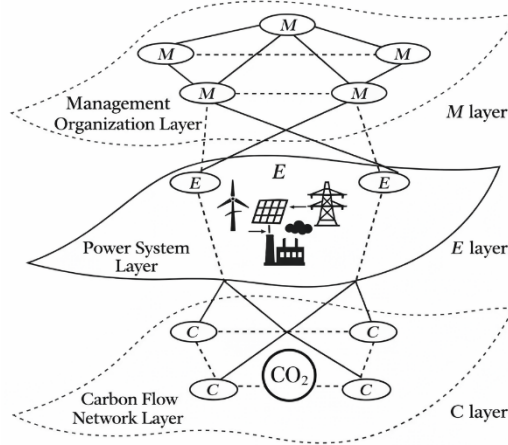
However, common shortcomings still exist, specifically: the three-dimensional separation remains a bottleneck. Both cases reflect that insufficient coupling between the “management organization layer (policy coordination, resource integration) - power system layer (source-grid-load-storage coordination) - carbon flow network layer (precise carbon emission control and constraint)” is the key factor restricting efficiency improvement; incentive compatibility needs to be strengthened: how to more effectively and rigidly transmit carbon costs to the micro-decision-making level of enterprises and design sufficiently strong economic incentives, especially in terms of reducing green electricity costs and increasing energy storage revenue, is a problem to stimulate enterprises' endogenous emission reduction motivation.

3. Construction of Three-dimensional Network Model Mechanism

3.1 Construction of A Collaborative Optimization Mechanism for the “Government-Electricity-Carbon” Three-dimensional Network

The synergistic optimization of the electricity-carbon system in high-energy-consumption industrial parks primarily involves three core dimensions: the power system (E), the carbon flow network (C), and the management organization (M). In current practice, E and C often independently optimize around “energy supply efficiency” and “emission accounting” respectively. The assessment and incentives at the M level fail to form a continuous and measurable transmission link. There is a lack of effective dynamic coupling among the three, resulting in the disconnection of prices, constraint signals and operational decisions, making it difficult to achieve the coordinated goal of “optimal energy allocation and precise carbon emission control”. For this reason, this paper builds a three-dimensional hypernetwork model (Figure 1) including the management organization layer (M layer), the power system layer (E layer) and the carbon flow network layer (C layer) for the park scenario, and designs a closed loop of “policy flow-energy flow-carbon flow”: realizes the continuous transmission of policy flow to the decision-making of the source network load storage with dynamic carbon strength and electricity and carbon price linkage, and the carbon flow formed by the operation results is fed back to the policy and scheduling side in real time to form an observable and calibrated coordination mechanism. The three-layer network model is shown in Figure 1 below:

Figure 1: Three-dimensional network collaborative optimization model of Management Organization Layer - Power System Layer - Carbon Flow Network Layer



3.1.1 Management Organizational Layer: Multi-subject Collaborative Mechanism

The three-dimensional collaborative optimization of “government-electricity-carbon” in high-energy-consuming parks is highly dependent on the effective transmission and integration of the management organization layer. The core subjects of the M level include government supervision (the Development and Reform Commission, the Department of Ecology and Environment), park operation (management committee), and the main body of enterprises (high-energy-consuming enterprises, energy service providers). In reality, policy fragmentation and inefficiency of resource integration are common. The former is manifested in the inconsistent caliber of policy goals, the dispersion of tools, and the mismatch of timing; the latter is manifested in the low efficiency of cross-departmental and cross-enterprise matching of data, funds, indicators and other factors between the main subjects. In order to identify and quantify the above constraints, this study constructs two core indicators at the M level: policy coordination P_c and resource integration efficiency C_d .

Because there are often differences in ambiguity, graduality and compatibility in the expression of policy objectives, in order to maintain the consistency of policy objectives, the policy synergy is calculated by fuzzy set theory, and the calculation formula is shown in (1):

$$P_c = \frac{\sum_{i=1}^n \cdot \sum_{j=1}^m \mu_{A_i}(x) \cap \mu_{B_j}(x)}{n \times m} \quad (1)$$

μ_{A_i}, μ_{B_j} , are the target membership functions of policy A and policy B, respectively. At the same time, this paper uses the “centrality” index C_d in social network analysis to measure the efficiency of resource integration, and calculates the normalized value of the sum of the inverses of the shortest paths between all nodes in the network, as shown in formula (2), so that the overall compactness of the network is more accurate.

$$C_d = \frac{\sum_{i=1}^N \cdot \sum_{j=1}^N d(i, j)^{-1}}{N(N-1)} \quad (2)$$

$d(i, j)$ is the shortest path length between nodes i and j and N is the total number of nodes. If the C_d value is greater than 0.7, the average path for information/resource transmission is shorter and more efficient, and the network is highly integrated; if the C_d value is less than 0.4, the average path is longer, the network will fragment, and the transmission of resources will be severely hindered.

3.1.2 Power System Layer: Source-Grid-Load-Storage Collaborative Model

As the physical carrier of the three-dimensional model, the E layer needs to realize the real-time linkage of energy flow dynamic optimization and carbon emission control. The model is built around the four types of objects of “power supply-power grid-load-energy storage”:

(1) Power supply side: Due to the timing dependence and meteorological sensitivity of photovoltaic output, this study integrates distributed photovoltaics, self-provided power plants and other multi-energy supply, and uses the LSTM-random forest combination model to predict the photovoltaic output to improve the prediction accuracy. The formula is shown in (3):

$$P_{pv,t} = f \left(\begin{matrix} LSYM(X_{t-n}, \dots, X_t) \\ RandomForest(T_t, G_t, V_t) \end{matrix} \right) \quad (3)$$

X_t is historical data, T_t is temperature, G_t is total horizontal irradiance, V_t is wind speed.

(2) Load side: Based on the principle of economy, this study also needs to consider the balance of scheduling flexibility and default cost. Therefore, a time-break response model of high-energy-consuming load $L_h(t)$ and adjustable load $L_a(t)$ is established on the load side, and the load interruption punishment item $P_p(t)$ is introduced. The punishment coefficient λ_t dynamically correlates the time-consuming electricity price with the carbon price, and drives the enterprise to actively adjust the adjustable load. The formula is shown in (4):

$$P_p(t) = \lambda_t \times (L_h(t) - L_{scheduled}(t)) \quad (4)$$

λ_t is the penalty coefficient for period t , $L_{scheduled}(t)$ is the rigid load demand.

(3) Energy storage side: Because the strong coupling relationship between charge and discharge power $P_{ch,d,t}$ and the charge state SoC_t will directly affect the stability of the system, the accurate modeling of the rated capacity E_r and efficiency η will be the prerequisite for ensuring the economy of energy storage. Based on the system characteristics of energy storage of lithium iron phosphate batteries in the energy storage system, this study constructs the charging and discharging state equation. The formula is shown in (5):

$$SoC_t = SoC_{t-1} + \eta \times \frac{P_{ch,d,t} \times \Delta t}{E_r} \quad (5)$$

E_t is the charge state at time t , η is the efficiency, $P_{ch,d,t}$ is the charging and discharging power, E_r is the rated capacity.

3.1.3 Carbon Flow Network Layer: Multi-dimensional Carbon Flow Accounting System

The core bottlenecks at the C layer of the electricity-carbon synergy in high-energy-consuming parks mainly lie in two aspects: First, the carbon cost accounting is static, with annual verification as the main approach, resulting in a mismatch between the scope, timeliness and operational decision-making. Second, the transmission of emission reduction targets is disconnected, and the target and price signals are difficult to be internalized into time-of-use scheduling. The industrial park frequently conducts annual carbon audits, which cannot respond to the real-time dispatching needs of the power system, resulting in a disconnect between the “electricity-carbon” system. To address this, this study constructs a three-tiered progressive mechanism based on the “production-consumption-trading” model, namely “accurate accounting → dynamic targets → endogenous incentives,” to streamline the transmission path from carbon constraints to energy decision-making.

The specific three-level collaborative mechanism is designed from three dimensions: carbon footprint accounting, carbon quota constraints, and carbon trading mechanisms.

In carbon footprint accounting, based on the characteristics of high-energy-consuming industrial parks that combine indirect emissions from electricity generation and emissions from industrial processes, this study sets localized factors according to the “Guidelines for Carbon Emission Accounting of Key Industries in Henan Province” to achieve provincial standard adaptation. Meanwhile, to avoid errors from a single method, this study adopts a hybrid accounting framework of “emission factor method + mass balance method”. The specific accounting formula is shown in (6):

$$C_t = \sum(E_t \times \gamma_e) + \sum(Min, t - Mout, t) \quad (6)$$

E_t is the activity data (electricity consumption), γ_e is the carbon emission factor of electricity, $Min, t/Mout, t$ is the carbon quantity of material input and output.

In terms of carbon quota constraints, in order to respond to the mandatory emission reduction requirements of the “Implementation Plan for Carbon Peaking in the Industrial Sector of Henan Province” and convert the static annual quota into a decreasing intensity threshold year by year, this study introduces a dynamic carbon intensity threshold calculation formula, as shown in (7):

$$C_{int,t} = C_{base} \times (1 - \alpha)^t \quad (6)$$

C_{base} is the reference carbon intensity, α is the annual reduction coefficient, t is the target year.

In terms of carbon trading mechanism, this study is linked with the market to build a cost optimization model under carbon price fluctuations, and uses the carbon price threshold (see (8) for the calculation formula) to trigger the coordinated scheduling of source network load storage. Reflect the product effect of quota gap and real-time carbon price in electricity consumption decision-making, and drive enterprises to achieve cost optimization through load adjustment.

$$C_{cost} = (E_t \times \gamma_e - Q_t) \times P_c(t) \quad (7)$$

E_t is electricity consumption, Q_t is carbon quotas, $P_c(t)$ carbon prices.

3.2 Data Platforms and Technical Tools

Data storage: Use MySQL to build a three-dimensional database, set up “policy management”, “energy flow” and “carbon flow” sub-tables, and realize cross-layer data association through a unique enterprise code (unified social credit code).

Analysis tools: Python (NLTK) is used for management data processing, MATLAB (optimization toolbox) is used for power and carbon data modeling, and Tableau is used for visualization.

3.3 Three-dimensional Network Coupling Optimization Model

The improved NSGA-II algorithm is adopted to solve the dual-objective optimization problem of economy and carbon emissions, and the objective function is:

$$\min(C_{total}, ECO_2) = \min(\sum t = 1T(C_{elec}, t + C_{ES}, t + C_{carbon}, t), \sum t = 1T E t C I t) \quad (9)$$

Among them, C_{elec}, t is the cost of time-of-use electricity pricing, C_{ES}, t is the operation and maintenance cost of energy storage, $C I t$ is the carbon intensity factor of the time period.

The following are the constraints:

Balance of power supply and demand:

$$P_{PV}, t + P_{coal}, t + P_{grid}, t = L_{high}, t + L_{flex}, t + P_{ES}, t \quad (10)$$

Among them, P_{PV}, t represents the photovoltaic power output in period t ; P_{coal}, t represents the coal-fired unit output in period t ; P_{grid}, t the grid purchase electricity in period t ; L_{high}, t represents the high-energy-consuming rigid load in period t ; L_{flex}, t represents the adjustable load in period t ; P_{ES}, t : the power of energy storage charging and discharging in period t (charging is negative, discharging is positive).

Energy storage operation boundary:

$$SoC_{min} \leq SoC_t \leq SoC_{max} \quad (11)$$

Among them, SoC_t represents the battery charge state (in percentage) of the energy storage at time period t ; SoC_{min} is the minimum allowable charge state of the energy storage; SoC_{max} is the maximum allowable charge state of the energy storage.

Carbon intensity constraint:

$$C I t \leq C I_{threshold}, t \quad (12)$$

Among them, $C I t$ represents the actual carbon intensity of period t ($tCO_2/100,000$ yuan of output value); $C I_{threshold}, t$ represents the carbon intensity threshold of period t ($tCO_2/100,000$ yuan of output value).

The model optimizes population diversity through the cross-entropy method, introduces elite retention strategies to improve the convergence speed, and finally generates the Pareto optimal solution, as shown in Table 1 below:

Table 1: Pareto final optimal solution set table

Total CO ₂ (Mt)	Unevenness (standard deviation)	Power proportion	Industry proportion	Ground Transport proportion
30.139886	15.23%	40.12%	31.64%	8.46%
27.585202	12.89%	43.85%	39.18%	9.81%
28.615207	14.02%	41.76%	35.92%	8.82%
26.821532	13.45%	43.21%	36.75%	9.12%
28.439863	14.56%	42.15%	38.01%	8.93%

4. Empirical Analysis and Results

4.1 Case Selection and Data Source

This study selects the national high-energy-consuming park of Petrochemical in Luoyang District, Henan Province as a typical research object. The industrial structure of the park covers typical high-energy-consuming industries such as electrolytic aluminum, steel and chemical industry. Its high-energy consumption and high-carbon emission characteristics are significantly representative in high-energy-consuming parks in Henan Province, which provides a solid foundation for the typicality and universality of the research conclusion. The time span of the research data is set to 2019-2024, and a comprehensive and timed data system is built.

In the data collection link, this study has established a systematic data architecture covering multiple dimensions and sources. The data at the management organization level mainly includes 89 policy documents issued by the park management committee. Through text mining and content analysis technology, the orientation and implementation path of policy formulation are deeply analyzed. At the same time, 120 carbon management questionnaires are distributed to enterprises in the park to obtain policy response and implementation from the enterprise practice level, and provide micro perspective support for policy coordination and resource integration efficiency analysis. The power system layer data covers the time-by-time electricity consumption data from 2019 to 2025. Relying on the smart meter system, more than 95% of the electricity consumption equipment in the park can be monitored in real time to achieve accurate capture of energy consumption dynamics; the actual measurement data of photovoltaic output has been calibrated by NASA POWER meteorological data, which significantly improves the accuracy of new energy power generation forecast; the dispatch data of Henan power grid sets boundary conditions for demand-side response research. In terms of carbon flow network layer data collection, based on the enterprise carbon verification report, the system collects information such as fossil energy consumption, industrial process emissions and indirect electricity emissions, and combines the quota price sequence of Henan carbon market in 2023-2024, enterprise carbon quota allocation and performance data, and real-time carbon emission data of the park's carbon monitoring platform to form a complete carbon flow tracking chain.

4.2 Data Acquisition and Processing

4.2.1 Multi-source Data Acquisition System

This study builds a three-dimensional data acquisition framework covering the “management organization layer-power system layer-carbon flow network layer” to ensure the cross-dimensional coupling and timeliness of the data.

(1) Manage organizational data:

Policy texts and management documents: 127 documents including the carbon quota allocation policy for parks from 2019 to 2025 and the green electricity trading rules were obtained through government public channels (the official website of the Development and Reform Commission of Henan Province and the government affairs platform of the Department of Ecology and Environment). Semi-structured interviews (n=23) were conducted with the Management Committee of the park, the Development and Reform

Commission, and the Department of Ecology and Environment to document the key nodes and implementation obstacles in policy transmission.

Enterprise questionnaire and synergy: distribute 120 carbon management questionnaires to high-energy-consuming enterprises (electrolytic aluminum, steel enterprises) to collect the response of enterprises to policies (photovoltaic installation willingness, energy storage investment decision-making); build a synergistic matrix of “policy makers-parks-enterprises” through social network analysis tools (UCINET) to quantify the policy synergy between nodes (the intensity of the relationship between carbon quota policies and electricity price policies).

(2) Power system layer data

Real-time data of energy flow: connect to the smart meter system of the park (covering more than 95% of electricity-consuming equipment), collect time-by-time electricity consumption data from 2019-2025 (resolution 15 minutes), including photovoltaic power station output (kW), self-prepared power plant power generation (MWh), energy storage system charging and discharging power (kW) and high-energy-consuming equipment load (such as electrolytic rectifier aluminum flow unit current/voltage data).

External environmental data: synchronously obtain NASA POWER meteorological data (including irradiance, wind speed and temperature) for photovoltaic power prediction model calibration; collect Henan power grid scheduling data (such as peak and valley period division, electricity price signals), and build boundary conditions for demand-side response.

(3) Carbon flow network layer data

Carbon footprint accounting data: Based on the enterprise carbon verification report (according to the “Guideline for Carbon Emission Accounting for Key Industries in Henan Province”), collect fossil energy consumption (coal and natural gas usage), industrial process emissions (such as blast furnace gas emissions of steel smelting) and indirect electricity emissions (electricity consumption $\times$ regional power grid carbon emission factor).

Carbon market and monitoring data: obtain the quota price sequence of Henan carbon trading market from 2023 to 2024 (28-52 yuan/ton), the carbon quota allocation and performance of enterprises; connect with the carbon monitoring platform of the park to collect real-time carbon emission data of key enterprises (through CEMS online monitoring equipment).

4.2.2 Data Processing and Integration Methods

(1) Data cleaning and preprocessing

Missing value filling: Considering the correlation of meteorological data, LSTM interpolation is used for photovoltaic output data; based on the co-occurrence matrix of policy keywords, the multiple interpolation method is adopted for the policy synergy degree data.

Abnormal value detection: use the 3σ criterion to identify abnormal fluctuations in the power load data (such as the instantaneous load increase caused by equipment failure), and verify and correct it in combination with the enterprise production log.

Digitalization of management data: extract keywords such as “carbon emission reduction target” and “subsidy strength” in the policy text through NLP technology to build a policy synergy index; use social network analysis to calculate the node center, and identify the “key intermediary node” in policy transmission.

Electric carbon data coupling: match the power flow data with the carbon flow data, and calculate the indirect carbon emissions of each electrical equipment through the current tracking algorithm to form a three-dimensional data table of “equipment-time-carbon intensity”.

(2) Data quality control

Reliability and validity testing: The policy interview data was verified through triangulation (cross-referencing documents and interview content) to ensure reliability; the electricity and carbon data were verified through equipment validation (such as regular calibration of electricity meters and comparative testing of carbon monitoring equipment) to control validity.

Collinearity handling: Perform VIF (variance inflation factor) tests on the input variables of the 3D network model, and use principal component analysis (PCA) to reduce highly correlated variables (such as electricity price and load) to a comprehensive index.

4.3 Analysis Results of Three-dimensional Network Nodes

According to data from Carbon Monitor, as shown in Figure 2, carbon emissions from the power industry in Henan Province exhibited significant seasonal fluctuations from 2019 to 2024, with double peaks in winter (January-February) and summer (June-August) due to heating and cooling demand, respectively. Notably, March and October saw emission troughs, possibly related to favorable weather and reduced electricity load. The Pareto optimal date (March 20, 2019) largely fell within these troughs, suggesting the potential value of seasonal adjustments for emission reduction.

Figure 2: The trend of daily CO₂ emissions of residents in Henan Province from 2019 to 2024

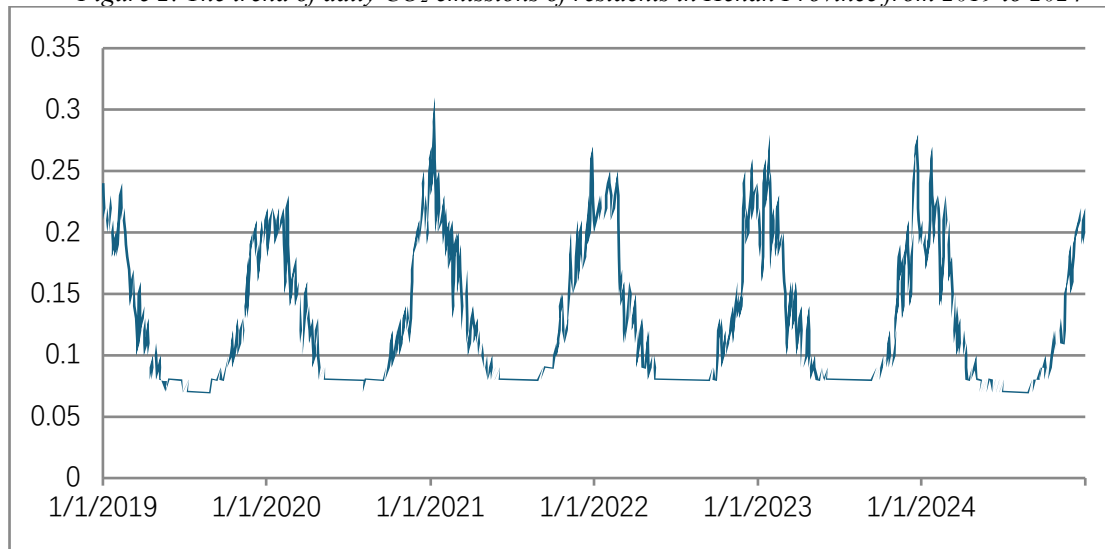


Figure 3: The daily variation trend of industrial carbon emissions in Henan Province from 2019 to 2024

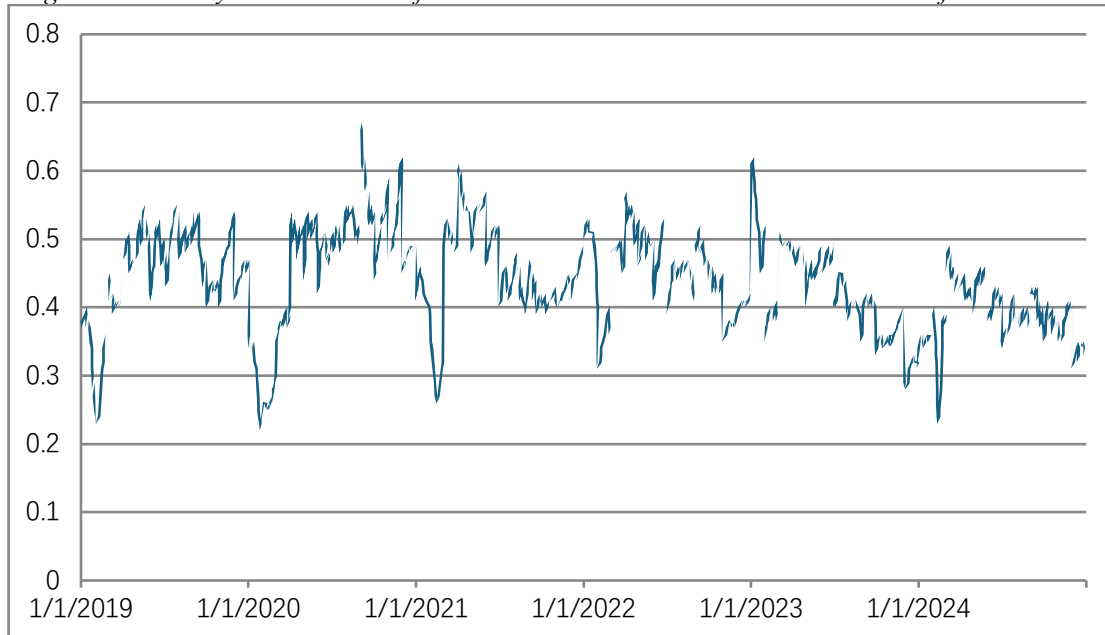
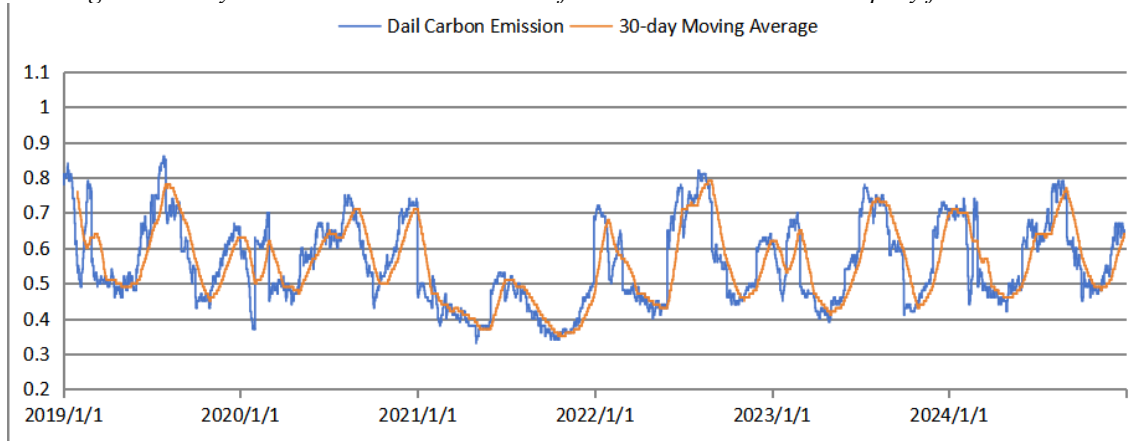


Figure 4: Daily carbon emission trend chart of Henan Electric Power Company from 2019 to 2024



4.3.1 Management Efficiency of the Organizational Layer

After quantitative analysis and measurement, the value of the policy synergy index at the management organization level P_c is 0.62, which shows that there is still significant room for optimization in the current policy system in terms of coordination mechanism. In-depth deconstruction found that the two core policies of “carbon quota allocation” and “green electricity trading” have obvious gaps in connection at the level of institutional design. The quantitative evaluation based on the semantic similarity algorithm shows that the relevance of the two policies is only 0.45, which significantly weakens the overall implementation effectiveness of the policy system. The resource integration efficiency index C_d is 0.78. Although it is at a medium level, the average path length of information transmission between the park management committee and enterprises, which is the center of policy transmission, reaches 2.3 jumps, and there are 1-2 key “structural holes” in the network topology. Such network structure defects seriously restrict the timeliness of policy transmission and the optimization process of resource allocation, resulting in the failure of the overall coordination function of the management organization level.

4.3.2 Optimization Effect At the Power System Layer

Under the three-dimensional collaborative optimization mechanism of “government-electricity-carbon”, the operation of the power system has been significantly improved in terms of economic benefits and system performance. From the analysis of economic benefits, the arbitrage income of Fenggu electricity price increased by 21.8% compared with the benchmark scenario, and the static investment recovery period of energy storage projects was shortened to 4.6 years, which effectively reduced the comprehensive energy cost of enterprises and provided a solid economic feasibility basis for investment decision-making of energy storage facilities. In terms of system performance, the consumption capacity of new energy has been significantly enhanced, the penetration rate of photovoltaic power generation has increased from 15% to 28%, the proportion of power supply load of self-provided power plants has decreased by 12 percentage points, and the peak and valley difference rate of the system has narrowed by 19%. The above optimization results show that the coordination mechanism has effectively promoted the low-carbon transformation of the power supply structure and realized the coordinated improvement of the stability and economy of the power system.

4.3.3 The Emission Reduction Achievements of the Carbon Flow Network Layer

The optimization results of the carbon flow network layer have been remarkable. The carbon intensity of the park has decreased from 1.85 tCO₂/10,000 yuan output value to 1.53 tCO₂/10,000 yuan output value, a decrease of 17.4%, which provides strong support for the low-carbon transformation of the park. Under the action of the carbon market mechanism, the carbon transaction cost is saved by 3.2 million yuan per year; the carbon emission factor per unit of electricity decreased from 0.82 kgCO₂/kWh to 0.68 kgCO₂/kWh, which realizes the coordinated improvement of environmental benefits and economic benefits, and verifies the feasibility of coordinated promotion of carbon emission reduction and economic development.

4.4 Verification of Coupled Model

The structural equation model is used to empirically test the coupling effect of the three-dimensional network of “management organization layer-power system layer-carbon flow network layer”. The results show that there is a significant influence relationship between the layers. As shown in Table 1:

Table 2: Estimation of three-dimensional network coupling effects

Pathway	Standardization coefficient	t value	P value
Management Organization Layer → Power System Layer	0.37	4.21	<0.01
Power System Layer → Carbon Flow Network Layer	0.52	5.83	<0.01
Carbon Flow Network Layer → Management Organization Layer	0.21*	2.15	<0.05

Note: * indicates that the indirect effect is conducted through the intermediary at the power system level

The direct effect of the management organization layer on the optimization of the power system layer is 0.37, and the scale of photovoltaic installation is improved through policy guidance; the path coefficient of the power system layer and the carbon flow network layer is 0.52, and the cleaning of power structure is the core driving force for carbon emission reduction; model fitting indicators: CFI=0.94, RMSEA=0.062, good suitability.

4.5 Model Compatibility Analysis

4.5.1 Multi-scenario Robustness Verification

To evaluate the adaptability of the model under different policy and technical conditions, three typical scenarios were designed for sensitivity testing. The results are shown in Table 3:

Table 3: Multi-scenario robustness test results

Scene type	Core parameter changes	Key indicator changes	Optimization measure
The carbon price has fluctuated sharply	The carbon price range is 28 to 42 yuan per ton	The standard deviation of total cost has expanded to 12.7%	Dynamically adjust the carbon quota allocation coefficient $\beta=0.85$
High penetration of new energy	The installed capacity of photovoltaic power has been doubled	The reduction in carbon intensity reached 22.3%	The energy storage capacity has increased by 18%
Extreme load fluctuation	The daily load rate has dropped from 0.65 to 0.52	The frequency of energy storage calls has increased by 41%	Load clustering compensation mechanism ($\lambda=1.15$)

4.5.2 Cross-regional Scalability Verification

Based on the research gap of regional differentiation pointed out in the literature review, the parks in Shandong (coastal) and Shaanxi (inland) were selected for horizontal comparison (Table 4), as shown in Table 4:

Table 4: Comparison of cross-regional model adaptation effects

Index	Henan (Benchmark)	Shandong Park	Shaanxi Park
Reduction in carbon intensity	17.4%	15.2%	19.1%
Policy synergy degree P_c	0.62	0.71 (+14.5%)	0.54 (-12.9%)
Model correction coefficient	-	0.92	1.08

From this, it is found that in terms of regional characteristics, the weight of energy storage allocation needs to be reduced by 8% due to the high green electricity penetration rate of 32% in Shandong Park, and the arbitrage income of Fenggu has increased to 25.3%. The result is consistent with the sensitivity analysis of the green electricity weight (QRES) in the carbon intensity accounting formula in Section 3.1.3, which can verify the amplification effect of the proportion of new energy on economic goals. Due to the low policy coordination of 0.54, Shaanxi Park needs to be improved by 23% through node centrality optimization to strengthen the driving effect of the management organization layer. This discovery echoes the governance needs of the “structural hole” problem in the social network analysis of Section 3.1.1, and can prove that low-synergy areas need to give priority to optimizing the policy transmission path.

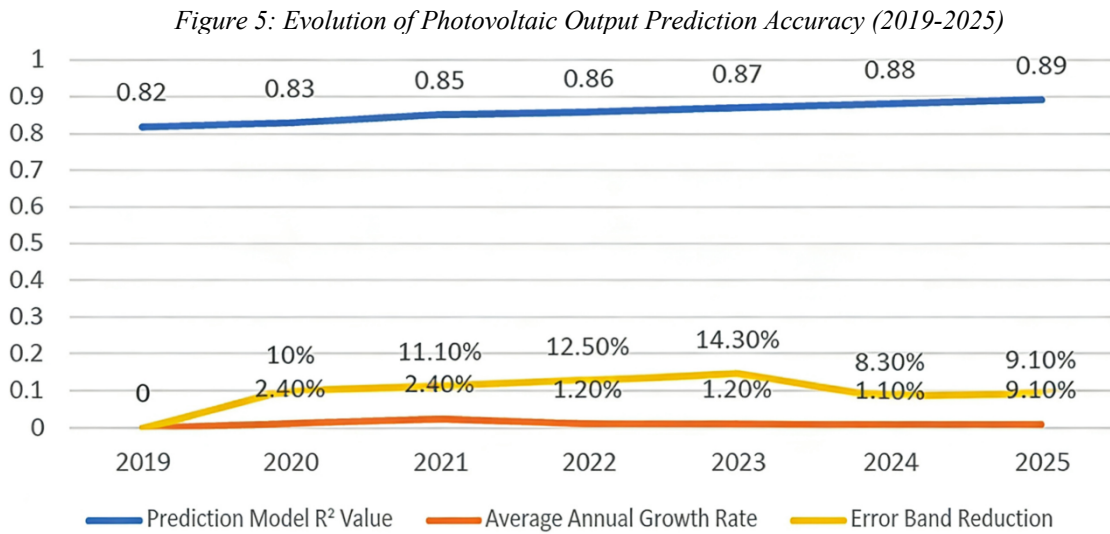
Therefore, this paper establishes a set of model adaptation correction mechanism to construct a regional characteristic index, as follows:

$$R = 0.35P_c + 0.45C_d + 0.2Q_{RES} \quad (13)$$

The parameter definition is inherited from the theoretical model in Section 3.1. The strong policy-driven mode is activated when $R > 0.7$, and the technology-first mode is enabled when $R < 0.5$. The differentiated performance of this mechanism in Shandong ($R=0.76$) and Shaanxi ($R=0.48$) can confirm the adjustability of the transmission path of “management organization layer → power system layer → carbon flow network layer” in the structural equation model of Section 4.3.

4.5.3 Analysis of Stability in the Time Dimension

Based on the rolling test of longitudinal data from 2019 to 2025, it was found that the MAPE of LSTM load forecast decreased from 4.7% to 3.1%, and the R^2 of photovoltaic output forecast increased from 0.82 to 0.89, as shown in Figure 5.



4.5.4 Technical Configuration Compatibility Testing

Based on the dynamic demand for the collaborative optimization of source network load storage in the three-dimensional network model, this study adopts multi-technical path alternative analysis, as shown in Table 5:

Table 5: Economic comparison of energy storage technologies

Technological type	Cost of investment (yuan/kWh)	Model optimization suggestions
Lithium iron phosphate battery	1500	Configure the capacity attenuation compensation mechanism ($\gamma=1.12$)
All-vanadium redox flow battery	3200	Requires subsidy ≥ 0.25 yuan/kWh
Compressed-air energy storage	4000	Applicable only for ≥ 100 MWh scene

Revealing the matching degree of energy storage technology characteristics and the power-carbon flow coordination scenario is the core basis for selection decision-making. Its key findings include: if lithium iron phosphate batteries are used, they perform best in terms of investment cost (1500 yuan/kWh) and response speed (climbing rate 5.2%/min), but a capacity decay compensation mechanism needs to be set (annual attenuation rate compensation coefficient $\gamma=1.12$). If all-vanadium liquid flow batteries are used, policy subsidies need to cover 38% of the initial investment difference (about 608 yuan/kWh) to achieve economic feasibility. If compressed air is used to store energy, due to the response speed limit (2.3%/min), the source load coordination efficiency will be reduced by 9%.

Based on all the above analysis, it can be seen that the three-dimensional network electric carbon optimization model shows strong adaptability under multi-scene, cross-regional and long-cycle conditions.

5. Conclusions and Implications

This article focuses on the reality of “increasing production means increasing carbon, and reducing carbon is afraid of loss” generally faced by high-energy-consuming parks. It builds a three-dimensional collaborative optimization mechanism of “government-electricity-carbon” covering the management organization layer, the power system layer and the carbon flow network layer, and conducts an empirical analysis of Luoyang Petrochemical High Energy Consumption Park in Henan Province as a case. Research shows that this mechanism can form a more effective linkage between policy transmission, energy scheduling and carbon emission control, providing a new analytical framework and realization path for high-energy-consuming parks to promote low-carbon transformation on the basis of ensuring production and operation.

Empirical results show that after the model was implemented, the carbon intensity per unit output value in the industrial park decreased from 1.85 tCO₂/10,000 yuan to 1.53 tCO₂/10,000 yuan, and the carbon emission factor per unit electricity decreased from 0.82 kgCO₂/kWh to 0.68 kgCO₂/kWh. Simultaneously, the photovoltaic penetration rate increased from 15% to 28%, and the static investment payback period for energy storage projects was shortened to 4.6 years. These results demonstrate that the three-dimensional synergistic mechanism has, to a certain extent, achieved simultaneous improvement in emission reduction and operational efficiency, indicating that the low-carbon transformation of high-energy-consuming industrial parks does not necessarily come at the expense of economic efficiency and production capacity.

The significance of this research lies in integrating policy mechanisms, power system operation, and carbon constraint management into a unified framework. This expands the research perspective on low-carbon governance in high-energy-consuming industrial parks and provides a reference for local governments and park managers to promote collaborative governance of energy transition and carbon reduction. In the future, it is possible to further combine the characteristics of different regional parks to deepen the research on policy coordination efficiency, market fluctuations impact, and engineering application feasibility, in order to enhance the universality and practical applicability of the model.

References

- [1] Department of Ecology and Environment of Henan Province. (2024). Bulletin on the eco-environmental status of Henan Province in 2023.
- [2] Department of Industry and Information Technology of Henan Province. (2023). Implementation plan for carbon peaking in the industrial sector of Henan Province (2023 edition) .
- [3] National Public Service Platform for Standards Information. (n.d.). Specification for circular economy evaluation of industrial parks (in Chinese).
- [4] National Development and Reform Commission. (n.d.). Guidelines for the green transformation of energy-intensive parks .
- [5] Chinese Government Network. (2023). Benchmark and baseline levels of energy efficiency in key industrial fields (2023 edition).
- [6] Yi, J., Zhao, M. X., Qin, X. H., Lin, W. F., Liu, C., & Deng, T. (2025). Exploration and practice of electricity-carbon synergy theory and technology. *Proceedings of the CSEE*, 45(19), 7427-7443.
- [7] Xiao, B., Zhang, Q. B., Qin, W. D., Qu, G. Q., Kong, Y. H., & Zhang, X. H. (2025). Collaborative planning method of source-grid-load-storage considering electricity-carbon coupling and flexibility supply-demand balance. *Electric Power Construction*, 46(3), 116-127.
- [8] Ji, X. Q., Li, Y. F., Ye, P. F., Yu, Z. H., Yang, M., & Zhang, Y. M. (2025). Power system planning with dual synergy of source-grid-load-storage and low-carbon economy. *Electric Power Automation Equipment*, 45(9), 174-182.

- [9] Zhang, G. L., Lin, H., & Tian, Y. Z. (2025). Source-load bi-level interactive optimization under the coupling of electricity, carbon and green certificate markets. *Power System Protection and Control*, 53(5), 1-10.
- [10] Dong, J. Z., Guo, H. Y., Jiang, T., Du, E. S., Zhang, N., & Kang, C. Q. (2025). Research on electricity-carbon coupling trading in the novel power system. *Proceedings of the CSEE*, 45(15), 5752-5771.
- [11] Li, X. S., Li, Z. Y., Qiu, C. L., Zhang, Y. H., Tao, Y. L., & Wang, Q. J. (2025). Operating strategy of electricity-carbon market for integrated energy system considering carbon market clearing. *Southern Grid Technology*, 19(4), 146-160.
- [12] Zhang, L. R., Yang, Z. F., & Cheng, C. Q. (2022). Emission reduction strategy selection of closed-loop supply chain under carbon quota trading policy. *Journal of Industrial Engineering and Engineering Management*, 36(1), 172-180..
- [13] Zhang, L. R., Liu, X. Y., Wang, F., Peng, B., & Lu, B. (2023). Research on misreporting decision and coordination of closed-loop supply chain under carbon quota trading policy. *Journal of Industrial Engineering and Engineering Management*, 37(4), 196-205..
- [14] Liu, P. D., Li, X. N., & Li, J. L. (2024). Low-carbon technology diffusion of competitive enterprises under the carbon quota trading mechanism: An evolutionary game analysis based on complex networks. *Systems Engineering - Theory & Practice*, 44(2), 684-703.
- [15] Li, B., Xia, X. Q., & Li, Q. Y. (2024). Impact of carbon trading on carbon emission reduction effect under carbon quota and its coordination mechanism. *Chinese Journal of Management Science*, 32(8), 250-261.
- [16] Ma, S., Liu, L., Cheng, H. Z., Zhang, X. H., Xu, L., & Lou, W. (2024). Bi-level flexible planning of source-grid-load-storage considering renewable energy and coal power based on scenario generation by adversarial neural networks. *Journal of Shanghai Jiao Tong University*, 1-22.
- [17] Cheng, Q. Y., Li, F., Luo, S. X., Wu, W. J., Zhang, N., & Kang, C. Q. (2025). Planning method framework and key supporting technologies for the novel power system. *Power System Technology*, 49(6), 2219-2231..
- [18] Liu, H., & Zhang, H. X. (2025). Hierarchical optimal scheduling strategy considering wind-solar-storage-carbon synergy and user satisfaction. *Automation of Electric Power Systems*, 49(11), 48-58.
- [19] Hong, M. Z., Wu, J. K., Luo, W. M., Sun, Y. H., Fang, Z. K., Zhan, Y. G., & Huang, Z. H. (2025). Carbon reduction collaborative optimization model for integrated energy systems considering cascade utilization of waste heat. *Southern Grid Technology*, 19(1), 129-139, 149.
- [20] Li, L. F., You, G. Z., Zhang, X., Wu, C., Cao, H. X., & Jin, Y. (2025). Research on optimal scheduling strategy for microgrid clusters in industrial parks considering electricity-carbon synergy. *Distribution & Utilization*, 42(3), 37-45, 67.
- [21] Huang, Y., Xue, Y. S., Xie, M., Huang, J., Shi, S. Y., & Yu, T. (2023). Decision optimization method for electricity-carbon collaborative trading of power generation groups under price-taker mode. *Automation of Electric Power Systems*, 47(21), 25-35.
- [22] Xu, Z. H., Chen, Z. P., Zeng, Z. H., Li, Z. K., Dong, G. Y., & Lin, J. K. (2025). Robust unit commitment model considering collaborative optimization of multiple reserve resources from source-grid-load-storage. *Proceedings of the CSU-EPSCA*, 37(1), 101-111.
- [23] Ding, Y. H., Tan, Q. L., Wei, Y. M., & Shan, Z. J. (2025). Collaborative optimization and operation of wind-solar-thermal power based on low-carbon sustainability under the coupling of electricity and carbon markets. *Systems Engineering - Theory & Practice*, 45(8), 2643-2656.
- [24] Chen, H. X., Kuang, S. N., Chen, X. Y., Li, X. Y., Zheng, X., Cheng, R., & Shi, L. (2023). Synergistic path of pollution reduction and carbon mitigation in wastewater treatment plants based on multi-objective optimization: A case study of a plant in Beijing. *Research of Environmental Sciences*, 36(11), 2148-2158..

- [25] Hou, J. Y., Chen, D. H., & Liu, W. W. (2022). Concept, model and optimization path of synergistic degree between pollution reduction and carbon mitigation: Based on municipal sludge data from 31 provinces and cities in China. *Construction Economy*, 43(S1), 566-569.
- [26] Zhang, G. L., Lin, H., & Tian, Y. Z. (2025). Source-load bi-level interactive optimization under the coupling of electricity, carbon and green certificate markets. *Power System Protection and Control*, 53(5), 1-10.
- [27] Zhou, F. H., Wang, H., Wang, H. L., Wang, M., Jin, Y. J., Li, C. C., ... & Wang, P. (2025). Research on multi-market subject behavior under electricity-carbon-green certificate coupled market based on multi-agent reinforcement learning. *Electric Power*, 58(4), 44-55.
- [28] Shao, Z. G., Lin, Y. Q., Chen, F. X., Zheng, X. H., Guo, Y. X., & Yan, X. Y. (2024). Distributed robust low-carbon economic dispatch of integrated electricity-gas-heat energy system considering carbon-green certificate coupling mechanism. *Electric Power Automation Equipment*, 44(6), 50-58, 76.

Funding

This research was supported by the Innovation and Entrepreneurship Training Program for College Students of Xi'an International Studies University, project number: 202510724038.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

An Analysis of Regional Differences in Consumption Expenditure Levels in China Based on Hierarchical Clustering

Yawen Shi*

School of Mathematics and Statistics, Beijing Technology and Business University, Beijing 100071, China

**Corresponding author: Yawen Shi.*

Abstract

There are significant regional disparities in the consumption expenditure levels of urban residents in China. Scientifically identifying the regional differentiation characteristics of consumption structure is of great significance for formulating differentiated regional development policies. Based on the data of per capita consumption expenditure of urban residents across eight major categories in 31 provinces, autonomous regions, and municipalities from the *China Statistical Yearbook 2024*, this study conducts an empirical analysis of regional differences in consumption expenditure levels in China using the systematic clustering method, specifically the within-cluster sum of squares method and squared Euclidean distance. First, descriptive statistics and boxplots are employed to identify the distribution characteristics of variables and detect outliers. Second, Pearson correlation coefficients are used to examine the relationships among variables. Subsequently, systematic clustering based on the within-cluster sum of squares method is performed, and the optimal number of clusters is determined with the aid of scree plots and dendrograms. Finally, one-way analysis of variance (ANOVA) is applied to test the statistical significance of the clustering results. The empirical results indicate that the 31 provinces in China can be classified into three consumption tiers. The first category is characterized as “high consumption–comprehensive development,” including Beijing, Shanghai, and Zhejiang, where the consumption structure has shifted toward development- and enjoyment-oriented consumption. The second category is defined as “moderate consumption–structural optimization,” covering Tianjin, Inner Mongolia, Jiangsu, Fujian, Guangdong, and Tibet, where the consumption structure is in a transitional stage of upgrading. The third category is described as “relatively low consumption–survival-oriented,” comprising the remaining 22 provinces, where consumption remains concentrated on essential goods. Significant inter-group differences are observed across all eight categories of consumption expenditure among the three groups ($p < 0.01$), and the clustering results are highly consistent with China’s regional economic development patterns. This study provides empirical evidence for understanding regional differences in consumption structure in China and for formulating differentiated consumption promotion policies.

Keywords

systematic clustering method, regional consumption disparity, consumption structure, urban residents, within-cluster sum of squares method

1. Introduction

1.1 Research Background

Consumption is the primary driver of economic growth and an important indicator for measuring high-quality regional economic development [1]. In recent years, the consumption levels of urban residents in China have significantly increased, with the consumption structure gradually shifting from survival-oriented to development- and enjoyment-oriented. However, influenced by multiple factors such as regional resource endowments, economic development levels, and institutional environment, there exist significant imbalances in the consumption expenditure levels of urban residents across different provinces in China [2]. Existing studies indicate a pronounced “polarization” in consumption tiers between the eastern and central-western regions, and the rapid increase in rigid expenditures such as housing and education has exerted a significant crowding-out effect on consumption in other regions, limiting the full potential of domestic demand [3, 4]. How to scientifically identify such regional differentiation characteristics has become a key issue to be addressed in formulating differentiated regional coordinated development strategies.

1.2 Research Questions

Although existing literature has paid attention to regional differences in consumption structure, most studies focus on single-variable analysis or a simple division into the eastern, central, and western regions, lacking fine-grained classification based on the latest multidimensional data. Specifically, current studies have not fully addressed the following questions: At the provincial level in China (31 provinces), based on per capita consumption expenditure data across eight major categories, what clustering patterns emerge among different regions? What are the essential statistical differences in consumption structure among these categories? These questions require rigorous empirical analysis to clarify and fill the gap in the precision of existing classification studies.

1.3 Research Objectives and Significance

This study aims to use the latest data from the *China Statistical Yearbook 2024* and apply the systematic clustering method to scientifically classify the consumption expenditure levels of urban residents in 31 provinces, thereby revealing the spatial differentiation patterns of regional consumption structure.

Theoretically, this research enriches the interdisciplinary perspective of regional economics and consumption economics and provides new empirical evidence for understanding the characteristic pattern of regional consumption gradients in China. Practically, the empirical results assist government authorities in accurately identifying the consumption development stages of different regions, enabling the formulation of more targeted consumption promotion policies. In particular, proposing differentiated strategies for “high consumption–comprehensive development” and “relatively low consumption–survival-oriented” regions has important practical significance for promoting a healthy development pattern of “high consumption leading, moderate consumption optimizing, and low consumption potential releasing” [5].

1.4 Research Methods

This study adopts quantitative empirical research methods. First, the per capita consumption expenditure of urban residents in eight major categories across 31 provinces, autonomous regions, and municipalities is selected as the research object. Second, descriptive statistics and boxplots are used to identify data distribution characteristics and detect outliers, and Pearson correlation coefficients are applied to examine the relationships among variables. In the core analysis stage, clustering is conducted using the systematic clustering method, specifically the within-cluster sum of squares method combined with squared Euclidean distance, and the optimal number of clusters is determined with the aid of scree plots and dendrograms. Finally, one-way analysis of variance (ANOVA) is employed to test the statistical significance of the clustering results, ensuring the scientific rigor and robustness of the classification.

1.5 Structure of the Paper

The paper is organized into five chapters. Chapter 1 is the introduction, outlining the research background, questions, objectives, methods, and framework. Chapter 2 describes the variable selection and data sources,

and conducts descriptive statistics and preliminary data processing. Chapter 3 presents the core empirical analysis, detailing the systematic clustering process, the optimal classification results, and the consumption structure characteristics of each category. Chapter 4 tests the statistical significance of the clustering results using one-way ANOVA and provides discussion. Chapter 5 concludes the study and offers policy recommendations, summarizing the main findings and proposing corresponding measures.

2. Variables and Data Sources

The data used in this study are sourced from the “Living Standards” section of the *China Statistical Yearbook 2024*, specifically the “Per Capita Consumption Expenditure of Urban Residents by Region.” The dataset covers per capita expenditures of urban residents in eight major consumption categories across 31 provinces, autonomous regions, and municipalities (excluding Hong Kong, Macao, and Taiwan). The variables include: food, tobacco, and alcohol; clothing; housing; household goods and services; transportation and communication; education, culture, and recreation; healthcare; and other goods and services.

All data correspond to actual expenditures in 2023 (unit: RMB). The data have been standardized to eliminate dimensional effects and facilitate subsequent clustering analysis.

3. Descriptive Statistical Analysis of Variables

3.1 Means and Standard Deviations of Variables

Before performing clustering analysis, descriptive statistics were conducted on the eight consumption expenditure categories, including sample size, minimum, maximum, mean, and standard deviation, to preliminarily understand the central tendency, dispersion, and distribution characteristics of regional consumption expenditures. The results are shown in Table 1.

Table 1: Means and Standard Deviations of Variables (Unit: RMB)

	N	Min	Max	Mean	SD
Food, tobacco, alcohol	31	6216	13568	9129.52	1802.764
Clothing	31	1029	2893	1880.61	368.897
Housing	31	4361	20261	7364.42	3790.403
Household goods & services	31	1286	2848	1818	383.375
Transportation & communication	31	3099	7558	4288.48	1011.633
Education, culture & recreation	31	1628	5339	3359.74	805.001
Healthcare	31	1551	4733	2873.32	694.477
Other goods & services	31	473	1849	929.32	338.834

From the perspective of mean values, expenditure on food, tobacco, and alcohol ranks first at 9,129.52 RMB, consistent with Engel’s law, indicating that food expenditure remains a core component of urban residents’ consumption structure, reflecting that China’s urban consumption is still dominated by basic living needs. Housing expenditure ranks second at 7,364.42 RMB, indicating that housing-related costs are another major pillar in urban residents’ consumption structure, highlighting the significant role of housing costs in daily spending. Expenditures on transportation and communication, education, culture and recreation, and healthcare are 4,288.48 RMB, 3,359.74 RMB, and 2,873.32 RMB, respectively, demonstrating that residents’ spending on travel, communication, human capital investment, and health protection has reached a certain scale, with development- and enjoyment-oriented consumption gradually emerging. In contrast, clothing, household goods and services, and other goods and services have relatively lower means of 1,880.61 RMB, 1,818.00 RMB, and 929.32 RMB, respectively, reflecting that these expenditures account for a relatively stable or still low proportion of total spending, with other goods and services exhibiting the highest elasticity, typically increasing significantly only after basic needs are met.

Regarding standard deviations reflecting dispersion, consumption variables exhibit significant regional differences. Housing expenditure has a standard deviation of 3,790.40 RMB, far exceeding other variables, with a maximum-to-minimum ratio of approximately 4.6, indicating substantial gaps in housing costs across regions. This phenomenon is closely related to differences in regional housing prices, rents, and property management fees and is a key factor causing imbalances in regional consumption expenditure levels.

Expenditure on food, tobacco, and alcohol also shows considerable variation ($SD = 1,802.76$ RMB), indicating that even basic living consumption differs significantly across regions, likely due to variations in prices, income levels, and dietary culture. Transportation and communication and education, culture, and recreation have moderate standard deviations (1,011.63 RMB and 805.00 RMB, respectively), indicating regional differences, with economically developed or geographically large regions investing more in transportation and communication and spending more on human capital and cultural consumption. Healthcare shows a relatively lower standard deviation (694.48 RMB), reflecting its status as a rigid expenditure, though it may still vary due to population structure and health insurance policies. Clothing, household goods and services, and other goods and services exhibit small standard deviations (368.90 RMB, 383.38 RMB, and 338.83 RMB, respectively), indicating strong convergence nationwide, high market liquidity, and limited influence from regional differences.

Overall, consumption variables show varying degrees of regional dispersion, with housing and food, tobacco, and alcohol exhibiting the most pronounced differences, fully reflecting the significant regional imbalance in urban residents' consumption structure. To further eliminate the effects of differences in units and magnitudes among variables and ensure the scientific and comparable nature of clustering results, the eight variables are standardized in subsequent analyses to construct a reasonable basis for regional classification.

3.2 Outlier Analysis of Variables

To examine the presence of outliers in the data and avoid their potential interference with subsequent clustering analysis, boxplots were used to identify outliers for the eight categories of consumption expenditure. Boxplots visually display the median, quartiles, extreme values, and possible outlier observations for each variable, aiding in assessing the distribution shape and the degree of deviation in the sample data.

Figures 1 to 8 present the boxplots for the eight categories of consumption expenditure. Overall, there are noticeable differences in the occurrence of outliers across variables among regions.

Figure 1: Boxplot of Food, Tobacco, and Alcohol

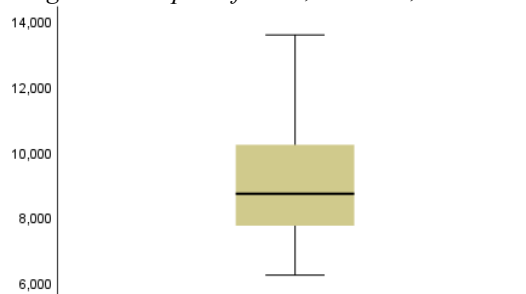


Figure 2: Boxplot of Clothing

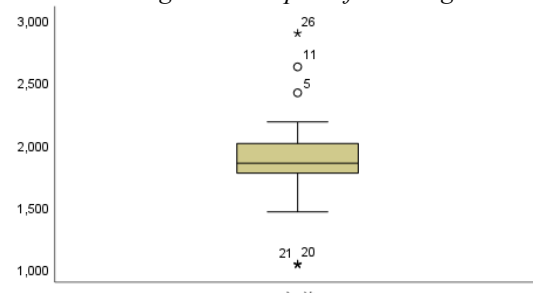


Figure 3: Boxplot of Housing

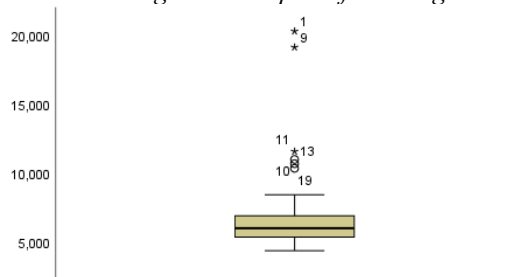


Figure 4: Boxplot of Household Goods and Services

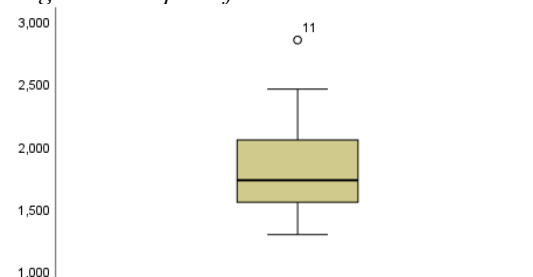
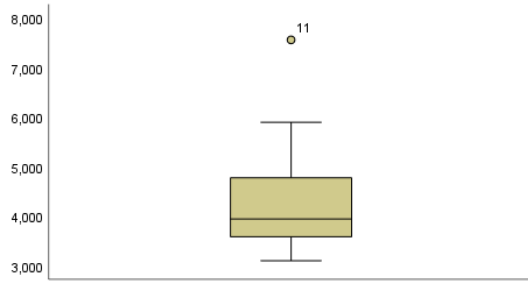
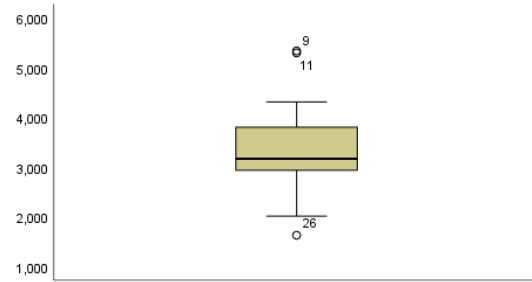
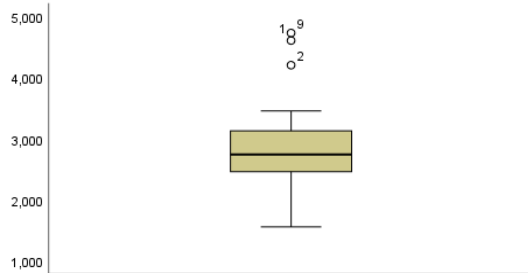
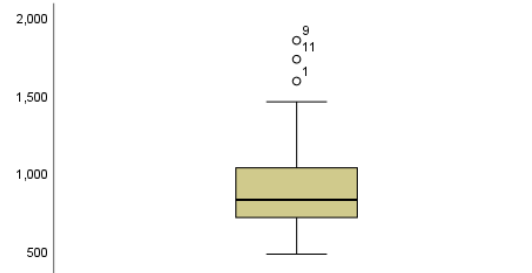


Figure 5: Boxplot of Transportation and Communication*Figure 6: Boxplot of Education, Culture, and Recreation**Figure 7: Boxplot of Healthcare**Figure 8: Boxplot of Other Goods and Services*

Based on the boxplot identification results, high-value outliers are observed in variables such as housing, food, tobacco and alcohol, transportation and communication, and education, culture, and recreation, particularly concentrated in economically developed eastern regions such as Beijing, Shanghai, Zhejiang, and Guangdong. The objective presence of outliers indicates significant regional imbalance in consumption structures, which may affect the measurement of inter-sample distances in clustering analysis. Since the complete linkage method is sensitive to outliers, the presence of extreme values may cause clustering results to be overly influenced by these points, potentially distorting the formation and classification of clusters. Therefore, in the choice of clustering methods and interpretation of results, caution should be exercised when using complete linkage, and results should be compared with other clustering approaches to enhance robustness and reliability.

3.3 Correlation Analysis of Variables

To examine the relationships among the eight consumption expenditure variables and prevent potential bias in subsequent clustering analysis due to multicollinearity, Pearson correlation coefficients were calculated to assess pairwise linear correlations among variables. The results are presented in Table 2.

As shown in Table 2, the consumption expenditure variables exhibit varying degrees of positive correlation, with most correlations being statistically significant. Specifically, strong positive correlations exist between housing and other goods & services ($r = 0.795$), transportation & communication and other goods & services ($r = 0.789$), and healthcare and other goods & services ($r = 0.745$), indicating that regions with higher housing costs and stronger transportation and communication demands also tend to spend more on other high-end and quality-oriented services. Moreover, significant positive correlations are observed between food, tobacco, and alcohol and housing ($r = 0.734$), household goods & services and transportation & communication ($r = 0.754$), and education, culture & recreation and transportation & communication ($r = 0.722$), reflecting interlinkages among different expenditure categories, where economically developed regions generally show higher spending across multiple consumption dimensions.

Notably, some variable pairs show weak or insignificant correlations, such as clothing and food, tobacco, and alcohol ($r = 0.285$, $P > 0.05$), clothing and education, culture & recreation ($r = 0.140$, $P > 0.05$), and healthcare and food, tobacco, and alcohol ($r = 0.300$, $P > 0.05$), indicating that these categories vary independently across regions and are influenced by different economic and social factors.

Table 2: Correlation Matrix of Variables

		Food, Tobacco, Alcohol	Clothing	Housing	Household Goods & Services	Transportation & Communication	Education, Culture & Recreation	Healthcare	Other Goods & Services
Food, Tobacco, Alcohol	Pearson correlation coefficient								
	Significance (two-tailed)	1	0.285	.734**	.700**	.675**	.653**	0.3	.684**
Clothing	Pearson correlation coefficient								
	Significance (two-tailed)		0.121	0	0	0	0	0.101	0
Housing	Pearson correlation coefficient								
	Significance (two-tailed)	0.285	1	0.275	.706**	.488**	0.14	0.211	.551**
Household Goods & Services	Pearson correlation coefficient								
	Significance (two-tailed)	0.121		0.134	0	0.005	0.454	0.254	0.001
Transportation & Communication	Pearson correlation coefficient								
	Significance (two-tailed)	.734**	0.275	1	.657**	.592**	.677**	.655**	.795**
Education, Culture & Recreation	Pearson correlation coefficient								
	Significance (two-tailed)	0	0.134		0	0	0	0	0
Healthcare	Pearson correlation coefficient								
	Significance (two-tailed)	.700**	.706**	.657**	1	.754**	.636**	.368*	.704**
Other Goods & Services	Pearson correlation coefficient								
	Significance (two-tailed)	0	0	0		0	0	0.042	0
Food, Tobacco, Alcohol	Pearson correlation coefficient								
	Significance (two-tailed)	.675**	.488**	.592**	.754**	1	.722**	.437*	.789**
Clothing	Pearson correlation coefficient								
	Significance (two-tailed)	0	0.005	0	0		0	0.014	0
Housing	Pearson correlation coefficient								
	Significance (two-tailed)	.653**	0.14	.677**	.636**	.722**	1	.593**	.653**
Household Goods & Services	Pearson correlation coefficient								
	Significance (two-tailed)	0	0.454	0	0	0		0	0
Transportation & Communication	Pearson correlation coefficient								
	Significance (two-tailed)	0.3	0.211	.655**	.368*	.437*	.593**	1	.745**
Education, Culture & Recreation	Pearson correlation coefficient								
	Significance (two-tailed)	0.101	0.254	0	0.042	0.014	0		0
Healthcare	Pearson correlation coefficient								
	Significance (two-tailed)	Food, Tobacco, Alcohol	Clothing	Housing	Household Goods & Services	Transportation & Communication	Education, Culture & Recreation	Healthcare	Other Goods & Services
	Significance (two-tailed)	.684**	.551**	.795**	.704**	.789**	.653**	.745**	1
Other Goods & Services	Pearson correlation coefficient								
	Significance (two-tailed)	0	0.001	0	0	0	0	0	0

***p < 0.01 (two-tailed), significant; *p < 0.05 (two-tailed), significant

Overall, although certain correlations exist, multicollinearity is not severe, and the variable structure retains sufficient discriminative power, making it suitable for clustering analysis to explore regional consumption patterns. Since systematic clustering is based on inter-sample distances, high correlations among variables may imply overlapping information, but the method's validity does not depend on variable independence. To fully capture comprehensive differences across multiple consumption dimensions, squared Euclidean distance is adopted as the distance measure in subsequent clustering analysis. This distance metric amplifies differences, enhances cluster discrimination, and more clearly reveals the spatial differentiation of consumption expenditure levels across regions in China.

4. Analysis of Regional Differences in Consumption Expenditure Levels in China

4.1 Introduction to the Systematic Clustering Method

The systematic clustering method is one of the most widely used clustering techniques. Its fundamental concept is an “agglomerative” bottom-up clustering strategy. Specifically, the clustering process proceeds as follows: First, each sample is considered as an independent cluster, forming the initial classification. Second, the distances between clusters are calculated, and the two closest clusters are merged into a new cluster, with distances recalculated between the new cluster and the remaining clusters. This merging and distance updating process is repeated iteratively, reducing the number of clusters by one at each step until all samples are ultimately combined into a single overarching cluster. The clustering process can be intuitively visualized using a dendrogram, and researchers can determine an appropriate number of clusters based on the hierarchical structure of the dendrogram and complementary methods such as scree plots.

Considering the characteristics of single linkage, complete linkage, average linkage, centroid method, and the within-cluster sum of squares (Ward) method, as well as the distributional features of the data, the Ward method was selected to avoid the influence of outliers while balancing intra-cluster homogeneity and inter-cluster heterogeneity. Squared Euclidean distance was adopted as the distance measure, as it amplifies differences between samples and improves cluster discriminability. To determine the optimal number of clusters, a comprehensive assessment combining the hierarchical structure of the dendrogram and the elbow point in the scree plot was employed, allowing scientific identification of regional patterns in consumption expenditure levels and providing a reliable basis for subsequent regional classification and policy analysis.

4.2 Scree Plot for Clustering Analysis

Clustering analysis was conducted using the Ward method with squared Euclidean distance. The dendrogram and scree plot were jointly used to determine the optimal number of clusters, enabling scientific identification of differences in consumption expenditure patterns across regions in China.

The elbow point of the agglomeration coefficient indicates a relatively optimal clustering. Observation of the scree plot shows that when the number of clusters decreases from four to three, the agglomeration coefficient exhibits a significant increase, suggesting that dividing the 31 provinces into three clusters is reasonable. This classification scheme effectively distinguishes different consumption patterns while maintaining simplicity and interpretability.

4.3 Clustering Analysis Results

Based on the Ward method and squared Euclidean distance, a systematic clustering analysis was performed on urban residents' consumption expenditure data across 31 provinces. The samples were ultimately divided into three clusters, as shown in Table 3.

Figure 9: Dendrogram of Clustering

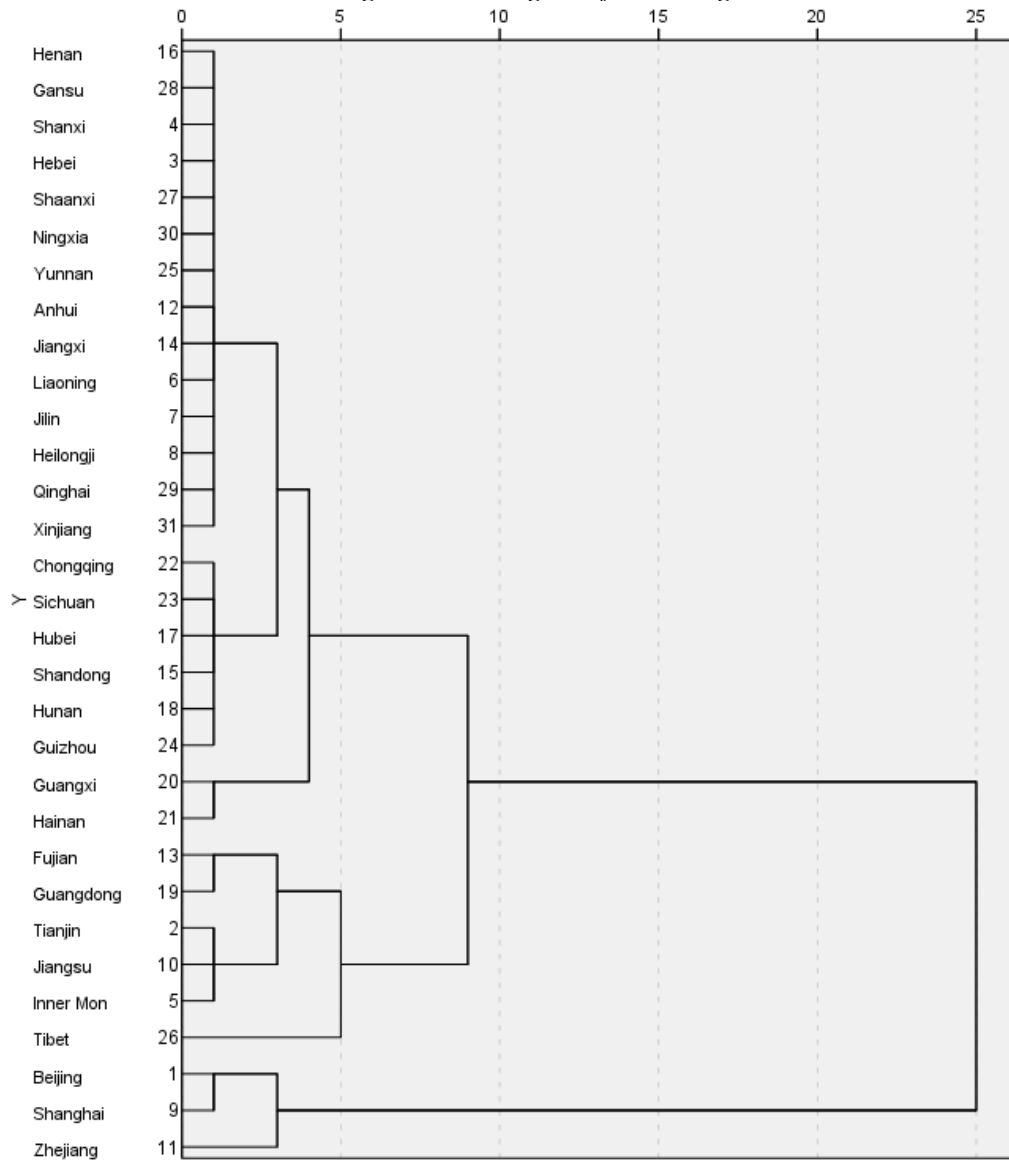


Figure 10: Scree Plot of Clustering Analysis

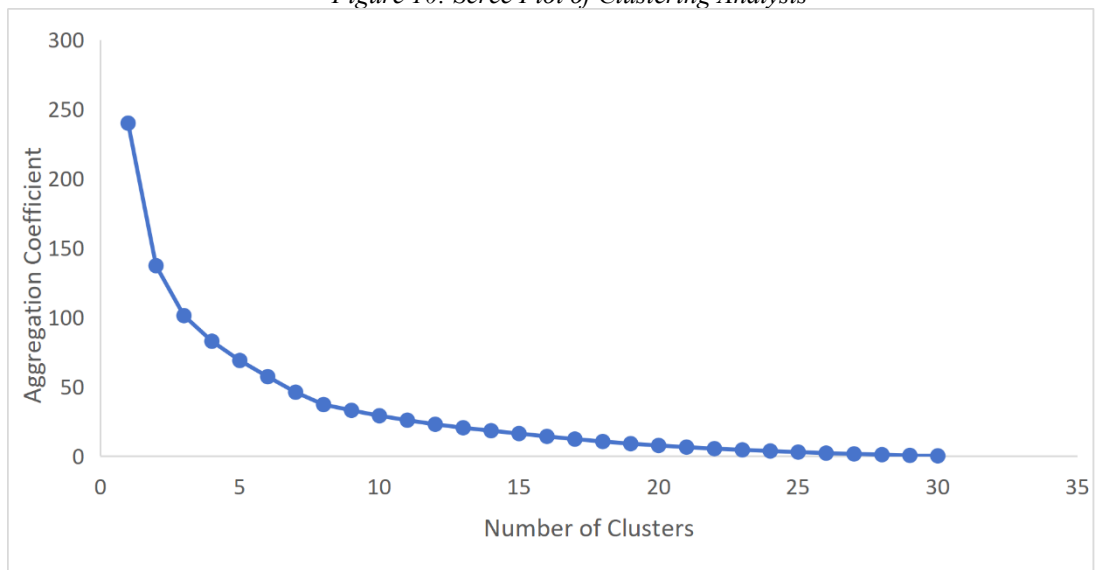


Table 3: Clustering Results

Cluster	Number of Provinces	Provinces
Cluster 1	3	Beijing, Shanghai, Zhejiang
Cluster 2	6	Tianjin, Inner Mongolia, Jiangsu, Fujian, Guangdong, Tibet
Cluster 3	22	Hebei, Shanxi, Liaoning, Jilin, Heilongjiang, Anhui, Jiangxi, Shandong, Henan, Hubei, Hunan, Guangxi, Hainan, Chongqing, Sichuan, Guizhou, Yunnan, Shaanxi, Gansu, Qinghai, Ningxia, Xinjiang

To further reveal the heterogeneity of consumption structures among clusters, Table 4 presents the within-cluster means and standard deviations for the eight consumption categories.

Table 4: Within-Cluster Means and Standard Deviations

Ward Method		Food, Tobacco, Alcohol	Clothing	Housing	Household Goods & Services	Transportation & Communication	Education, Culture & Recreation	Education, Culture & Recreation	Other Goods & Services
1	Mean	12353.67	2293.33	16970.33	2553	6175	4929.33	4196.67	1721.33
	Average	1567.597	290.107	4729.904	260.04	1266.432	674.368	823.906	131.127
2	Mean	10643.67	2136.67	8730.17	2086.83	5052.67	3469	2773.83	1133.5
	Average	1274.912	471.048	2298.228	152.089	780.176	953.398	918.727	242.05
3	Mean	8276.91	1754.05	5682.05	1644.45	3822.82	3115.91	2719.09	765.64
	Average	1058.368	271.542	675.442	261.997	511.753	498.376	381.887	137.236

From the group mean values reported in Table 4, it can be observed that the three clusters exhibit clear gradient differences in both consumption expenditure levels and consumption structure. A detailed analysis is as follows.

Cluster 1 represents high-consumption, comprehensive development regions. This cluster includes Beijing, Shanghai, and Zhejiang, which are among the most economically developed core regions in China. In terms of consumption levels, this group ranks at the top across all eight categories of consumption expenditure. In particular, expenditures on housing (16,970.33 yuan), transportation and communication (6,175.00 yuan), education, culture and entertainment (4,929.33 yuan), and healthcare (4,196.67 yuan) are especially prominent. These development- and enjoyment-oriented expenditures are significantly higher than those of the other two clusters. The consumption structure of urban residents in this cluster has largely moved beyond a survival-oriented pattern and has fully shifted toward development- and enjoyment-oriented consumption, representing the forefront of consumption upgrading in China. These regions typically feature higher income levels, well-developed infrastructure, and comprehensive public service systems, with consumption behavior reflecting typical characteristics of a post-industrial society.

Cluster 2 corresponds to moderate-consumption, structurally optimized regions. This cluster includes Tianjin, Inner Mongolia, Jiangsu, Fujian, Guangdong, and Tibet. In terms of composition, it encompasses both economically developed coastal provinces (Jiangsu, Fujian, Guangdong) and regions with distinctive policy or geographic conditions (Tianjin, Inner Mongolia, Tibet). Although there are differences in economic development levels, the consumption patterns are relatively similar. In terms of consumption levels, expenditures in this cluster fall within the national mid-range. Basic consumption categories such as food, tobacco and alcohol (10,643.67 yuan) and housing (8,730.17 yuan) are notably higher than those in Cluster 3, but there remains a substantial gap compared with Cluster 1 in development- and enjoyment-oriented consumption. It is noteworthy that the inclusion of Tibet suggests that consumption structure may be influenced by multiple factors, including policy support, demographic structure, and regional characteristics. Overall, Cluster 2 is at a critical stage of transformation and upgrading in consumption structure, exhibiting both features of advanced consumption patterns and constraints associated with regional development disparities.

Cluster 3 represents lower-consumption, survival-dominated regions. This cluster includes 22 provinces, mainly located in central, western, and northeastern China, where the level of economic development is relatively lagging and constitutes a major component of regional imbalance in China. In terms of consumption structure, this group remains at a generally low level across all eight categories of consumption expenditure. Consumption is still concentrated in essential categories such as food, tobacco and alcohol (8,276.91 yuan),

while the proportion of development- and enjoyment-oriented consumption is relatively limited. The consumption pattern in this cluster exhibits typical survival-oriented characteristics, with disposable income constraints limiting consumption upgrading. At the same time, this cluster covers a wide range of regions with considerable internal heterogeneity, and the consumption potential of some provinces has not yet been fully realized. From the perspective of expanding domestic demand and promoting coordinated regional development, this cluster represents an important potential space for future consumption growth and a key target for policy intervention and structural adjustment.

Overall, the classification of the three clusters reveals a clear gradient progression in both consumption levels and consumption structure, which fully validates the effectiveness of the clustering results. Cluster 1 shows a highly diversified consumption structure, Cluster 2 exhibits transitional characteristics of structural optimization, and Cluster 3 remains dominated by survival-oriented consumption. The differences among the three clusters not only reflect the imbalance in economic development levels but also reveal the spatial differentiation of consumption structure among urban residents in China, providing empirical evidence for the formulation of differentiated regional consumption policies.

4.4 One-way Analysis of Variance

To further verify the rationality and validity of the clustering results, a one-way analysis of variance (ANOVA) is employed to test the significance of differences among the three clusters across the eight categories of consumption expenditure. ANOVA determines whether statistically significant differences exist among groups by comparing the relative magnitude of between-group mean squares and within-group mean squares. If the between-group differences are significantly larger than the within-group differences, the clustering results can be considered to have strong discriminative validity. The results are presented in Table 5.

Table 5: One-way ANOVA Results

		Sum of Squares	df	Mean Square	F	Significance
Food, Tobacco, Alcohol	Between groups	60934003.92	2	30467001.96	23.331	0
	Within groups	36564741.82	28	1305883.636		
	Total	97498745.74	30			
Clothing	Between groups	1256817.433	2	628408.716	6.226	0.006
	Within groups	2826196.955	28	100935.606		
	Total	4083014.387	30			
Housing	Between groups	350280747.1	2	175140373.5	60.742	0
	Within groups	80733908.46	28	2883353.873		
	Total	431014655.5	30			
Household Goods & Services	Between groups	2716899.712	2	1358449.856	22.475	0
	Within groups	1692384.288	28	60442.296		
	Total	4409284	30			
Transportation & Communication	Between groups	18951261.14	2	9475630.568	22.579	0
	Within groups	11750776.61	28	419670.593		
	Total	30702037.74	30			
Education, Culture & Recreation	Between groups	8770473.451	2	4385236.725	11.507	0
	Within groups	10670340.49	28	381083.589		
	Total	19440813.94	30			
Education, Culture & Recreation	Between groups	5836417.456	2	2918208.728	9.457	0.001
	Within groups	8640529.318	28	308590.333		
	Total	14476946.77	30			
Other Goods & Services	Between groups	2721423.517	2	1360711.758	52.709	0
	Within groups	722839.258	28	25815.688		
	Total	3444262.774	30			

From the results in Table 5, it can be seen that at the significance level of $\alpha = 0.01$, the F-statistics for all eight categories of consumption expenditure pass the significance test, with corresponding p-values less than 0.01. Among them, variables such as housing and other goods and services exhibit relatively high F-values, indicating particularly strong between-group differences. Although the F-value for clothing is relatively smaller, its P-value is 0.006, which still indicates a statistically significant difference at the 0.01 level.

These results indicate that the three clusters derived from hierarchical clustering exhibit significant differences in group means across all eight consumption expenditure categories. In other words, the consumption structures of different clusters are not only characterized by overall gradient differences but are also statistically distinguishable across specific consumption categories. This provides statistical evidence supporting the rationality and validity of the clustering results and demonstrates that the classification effectively captures the true structure of regional differences in consumption expenditure levels in China, thereby providing a reliable basis for subsequent regional comparisons and policy analysis.

5. Conclusion

Based on data from the *China Statistical Yearbook 2024*, covering per capita consumption expenditures of urban residents across eight categories in 31 provinces, autonomous regions, and municipalities, this study applies hierarchical clustering using Ward's method and squared Euclidean distance to classify regional differences in consumption expenditure levels in China. Through descriptive statistical analysis, boxplot-based outlier detection, correlation analysis, hierarchical clustering, and one-way ANOVA, the following main conclusions are obtained.

First, there are significant regional disparities in the consumption structure of urban residents in China, which can be clearly classified into three categories with different consumption levels. The clustering analysis divides the 31 provinces into three groups: Cluster 1 includes Beijing, Shanghai, and Zhejiang, representing high-consumption, comprehensive development regions; Cluster 2 includes Tianjin, Inner Mongolia, Jiangsu, Fujian, Guangdong, and Tibet, representing moderate-consumption, structurally optimized regions; and Cluster 3 includes the remaining 22 provinces, representing lower-consumption, survival-dominated regions. This classification is validated by both the hierarchical structure of the dendrogram and the elbow point observed in the scree plot, indicating strong internal stability of the clustering scheme.

Second, the differences among the three clusters are reflected not only in the gradient distribution of total consumption expenditure but also more fundamentally in the internal composition of consumption structure. Cluster 1 exhibits significantly higher expenditures in development- and enjoyment-oriented categories such as housing, transportation and communication, education, culture and entertainment, and healthcare, with a consumption structure characteristic of a post-industrial society, where survival-oriented consumption declines and service- and development-oriented consumption becomes dominant. Cluster 2 is in a transitional stage of structural upgrading, with relatively balanced expenditures, reflecting both a trend toward consumption upgrading and a noticeable gap compared with Cluster 1 in higher-level consumption. Cluster 3 remains concentrated on essential consumption such as food, tobacco and alcohol, with a consumption pattern focused on meeting basic needs, a relatively low share of development- and enjoyment-oriented consumption, and underdeveloped consumption potential. The "high-medium-low" gradient pattern across the three clusters systematically reflects the regional differentiation of consumption structure in China.

Third, the clustering results demonstrate strong statistical significance and practical interpretability and are highly consistent with China's regional economic development pattern. The results of one-way ANOVA show that the mean differences in all eight consumption expenditure categories among the three clusters are statistically significant ($p < 0.01$), indicating strong discriminative validity of the classification. In terms of geographical distribution, Clusters 1 and 2 are mainly concentrated in the eastern coastal regions and economically developed or policy-advantaged areas, while Cluster 3 is widely distributed in central, western, and northeastern regions. Consumption levels exhibit a gradient decline from the eastern coastal areas to the inland regions. This distribution is highly consistent with the overall pattern of regional economic development in China and reveals a significant positive relationship between urban residents' consumption levels and regional economic development.

In summary, the empirical analysis of regional differences in consumption expenditure levels based on hierarchical clustering is methodologically rigorous and provides strong practical interpretability. The findings offer a reference for identifying regional consumption structure characteristics, formulating differentiated consumption promotion policies, and promoting coordinated regional development. Future research may further incorporate panel data to examine the dynamic evolution of consumption structure and introduce additional socioeconomic variables to explore the mechanisms underlying regional consumption disparities and potential policy interventions.

References

- [1] Shi, M., Jiang, Z., & Zhou, X. (2019). Consumption upgrading or consumption downgrading? *China Industrial Economics*, (7), 42–60. <https://doi.org/10.19581/j.cnki.ciejournal.2019.07.003>
- [2] Fu, H. (2020). Regional differences in residents' consumption structure in China: An empirical analysis based on the ELES model. *Consumer Economics*, 36(4), 77–84.
- [3] Tian, Q., Ma, J., & Gao, T. (2008). Regional differences in factors affecting urban residents' consumption in China. *Management World*, (7), 27–33. <https://doi.org/10.19744/j.cnki.11-1235/f.2008.07.004>
- [4] Tang, Q., Xia, Q., & Li, S. (2018). Consumption structure of urban households in China: 1995–2013. *Economic Research Journal*, 53(2), 35–49.
- [5] Zang, X., & Sun, W. (2003). Consumption structure of urban and rural residents: A comparative analysis based on the ELES and AIDS models. *Journal of Shandong University (Philosophy and Social Sciences)*, (6), 122–126.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

The Impact of the Negative List System for Market Access on Real Earnings Management: Evidence from a Multi-period DID Quasi-natural Experiment

Junhan Sun*

School of Economics and Management, Northeast Normal University, Changchun 130000, China

**Corresponding author: Junhan Sun.*

Abstract

This paper exploits the phased pilot implementation of the *Negative List System for Market Access* as a quasi-natural experiment. Using data on A-share listed firms in the Shanghai and Shenzhen stock exchanges from 2013 to 2019, we employ a multi-period difference-in-differences (DID) approach to examine the impact of this system on firms' real earnings management. The results show that the Negative List System for Market Access significantly reduces the level of real earnings management. This finding remains robust after a series of robustness checks, including propensity score matching (PSM) and entropy balancing. Heterogeneity analysis indicates that the policy effect is more pronounced in firms with larger board size, while it is relatively weakened in firms with CEO duality, abundant cash flows, and strong profitability. From the perspective of corporate financial behavior, this study reveals the micro-level governance effects of market access reform and provides empirical evidence for deepening the “streamlining administration, delegating power, strengthening regulation, and improving services” reform and optimizing the business environment.

Keywords

negative list system for market access, real earnings management, multi-period difference-in-differences, competition governance

1. Introduction

In the context of high-quality development, improving the information quality of the capital market has become an important prerequisite for promoting the efficient allocation of resources. The report of the 20th National Congress of the Communist Party of China explicitly proposes to “build a unified national market, deepen market-oriented reforms of factor markets, and develop a high-standard market system,” which provides clear guidance for market-oriented reforms in the new era. As an important institutional innovation for deepening the reform of “streamlining administration, delegating power, strengthening regulation, and improving services” and optimizing the business environment, the Negative List System for Market Access establishes the principle of “entry unless prohibited,” clearly specifying sectors that are prohibited or restricted, substantially reducing administrative approval items, lowering barriers to market entry, and promoting a fair and competitive market environment. Since 2015, this system has been implemented in batches across selected regions and was fully rolled out nationwide in 2018. Its gradual implementation process provides an ideal

quasi-natural experimental setting for evaluating the micro-level governance effects of market-oriented reforms.

Corporate earnings management behavior is a key factor affecting the information quality of the capital market. Compared with accrual-based earnings management, which relies on changes in accounting estimates, real earnings management influences accounting profits through adjustments in actual operating activities, such as altering production scale, manipulating sales policies, and cutting R&D expenditures. It is more concealed, harder to detect, and more detrimental to firms' long-term value. A large body of literature shows that real earnings management not only reduces financial reporting transparency and distorts resource allocation signals, but may also impair firms' core competitiveness, ultimately hindering high-quality economic development [1, 2]. Therefore, how to effectively constrain real earnings management from an institutional perspective has become an important issue in corporate governance and capital market research.

The implementation of the Negative List System for Market Access implies that industries previously subject to restrictions or implicit barriers are opened to more market participants. The entry of new firms enhances the adequacy and fairness of market competition. As an important external governance mechanism, whether market competition can be effectively transmitted to firm-level behavior and thereby constrain managerial incentives for earnings management remains an open question. Theoretically, the Negative List System for Market Access may influence corporate earnings management through the following channels: on the one hand, lowering entry barriers attracts more market participants, increases industry competition, and strengthens market discipline; on the other hand, improved institutional transparency helps regulate government behavior, optimize the business environment, and enhance external supervision. The combined effects of these channels are expected to restrain corporate earnings management behavior.

The literature related to this study mainly focuses on two strands. The first concerns the economic consequences of market access reform, where existing studies primarily examine firm investment [3], innovation activities [4], and resource allocation efficiency [5]. The second strand investigates the determinants of earnings management, with prior studies focusing on corporate governance characteristics [6] and external monitoring mechanisms [7]. However, few studies link market access reform with corporate earnings management or examine the impact of institutional changes on corporate financial behavior from the perspective of competition as an external governance mechanism. This paper attempts to fill this research gap.

Based on this, this paper uses the phased pilot implementation of the Negative List System for Market Access as a quasi-natural experiment. Using data on A-share listed firms from 2013 to 2019 and employing a multi-period DID approach, we systematically examine the impact of this system on real earnings management. The results indicate that the system significantly reduces firms' real earnings management, and this conclusion remains robust after a series of tests, including PSM-DID, entropy balancing, parallel trend tests, and placebo tests. Heterogeneity analysis shows that the policy effect is more pronounced in firms with larger boards, while it is relatively weaker in firms with CEO duality, abundant cash flows, and strong profitability.

The contributions of this paper are threefold. First, it extends the literature on the economic consequences of market access reform from the perspective of corporate financial behavior. Unlike existing studies that focus on investment and innovation, this paper examines earnings management as an important financial decision, revealing the micro-level governance effects of the Negative List System for Market Access. Second, it enriches the literature on the relationship between market competition and earnings management. By exploiting a quasi-natural experiment, this paper provides cleaner empirical evidence on how changes in the competitive environment affect corporate financial behavior. Third, it highlights the role of institutional environment in improving accounting information quality, offering policy implications for deepening market access reform, optimizing the business environment, and strengthening the foundational role of competition policy.

2. Institutional Background and Theoretical Hypotheses

2.1 Institutional Background

The market access system is an important institutional arrangement that regulates the entry of market participants into specific industries. For a long time, some industries in China have been characterized by high administrative entry barriers, requiring firms to undergo complex approval procedures to enter the market,

which has objectively restricted market competition and affected resource allocation efficiency. To further stimulate market vitality and optimize the business environment, China began to explore and implement the Negative List System for Market Access in 2015. By explicitly specifying industries that are prohibited or restricted, the system allows all types of market participants to enter sectors outside the list on an equal and lawful basis, thereby establishing the principle of “entry unless prohibited.”

In December 2015, the first batch of pilot programs was launched in Shanghai, Guangdong, Tianjin, and Fujian. In 2016, the pilot scope was expanded to 11 additional provinces, including Liaoning, Jilin, Heilongjiang, Zhejiang, Henan, Hubei, Hunan, Chongqing, Sichuan, Guizhou, and Shaanxi. In 2017, the pilot program was further deepened. In December 2018, the National Development and Reform Commission and the Ministry of Commerce jointly issued the *Negative List for Market Access (2018 Edition)*, marking the full nationwide implementation of the system. Compared with the traditional approval system, the Negative List System for Market Access reduces administrative approval items and improves the transparency of market rules, thereby effectively lowering market entry barriers and promoting the entry of more market participants. This regionally and temporally staggered pilot process provides substantial institutional variation for identifying policy effects using a multi-period DID approach.

2.2 Literature Review

Corporate earnings management has long been a central topic in accounting and corporate governance research. Existing studies generally classify earnings management into accrual-based earnings management and real earnings management [1]. Compared with accrual-based earnings management, which relies on accounting estimates and policy adjustments, real earnings management affects accounting profits by altering firms’ actual operating decisions, such as abnormal production, price discounts, and reductions in R&D and advertising expenditures. It is more concealed and harder to detect. Prior studies show that real earnings management not only reduces financial reporting quality but may also damage firms’ long-term value and lead to inefficient resource allocation [2].

Regarding the determinants of earnings management, existing literature mainly examines corporate internal governance, external monitoring mechanisms, and institutional environment. In terms of internal governance, factors such as board structure, the proportion of independent directors, and managerial power have been shown to significantly affect earnings management [8, 9]. In terms of external monitoring, institutional investors, analyst coverage, media scrutiny, and audit quality have been found to effectively constrain earnings manipulation [7, 10]. In recent years, with the rise of institutional economics, increasing attention has been paid to the impact of institutional environment on corporate financial behavior. Relevant studies indicate that a sound institutional environment can effectively restrain opportunistic behavior by improving market rules, strengthening property rights protection, and enhancing market transparency.

As an important institutional arrangement for deepening market-oriented reform, the Negative List System for Market Access has attracted increasing academic attention in recent years. Existing studies mainly examine its economic effects at the macro and meso levels, finding that it promotes firm investment and innovation, improves resource allocation efficiency, and enhances the local business environment. For example, Wang Xiongyuan and Xu Jing (2022) [3] find that the system significantly improves investment efficiency, mainly through enhanced competition and fairness as well as reduced transaction costs. Zhang Han et al. (2021) [5] show that the relaxation of market access regulation helps promote supply-side capacity reduction and improve capacity utilization. However, studies examining the impact of the Negative List System for Market Access from the perspective of corporate financial behavior, especially real earnings management, remain limited. The mechanism through which market competition, as an external governance mechanism, affects corporate financial behavior has not been fully explored.

2.3 Research Hypothesis

Real earnings management refers to firms’ manipulation of accounting profits through changes in actual operating decisions, such as abnormal production, early revenue recognition, or reductions in R&D expenditures. Compared with accrual-based earnings management, real earnings management is more concealed but has more profound negative effects on firms’ long-term value.

The implementation of the Negative List System for Market Access improves firms' external operating environment by lowering industry entry barriers and enhancing the transparency of market rules. Theoretically, this system may influence real earnings management through two channels. First, lower entry barriers attract more market participants, intensify market competition, and strengthen market discipline. In a more competitive environment, firms face greater operational pressure and higher risks of market exit, making them more reliant on real operational performance to maintain their market position, thereby reducing the scope for manipulating profits through real earnings management [11]. Second, increased institutional transparency helps regulate government approval processes, reduce implicit barriers, and enhance supervision by investors and regulators. Stronger external monitoring increases the likelihood and cost of detecting earnings management, thereby suppressing managerial incentives for manipulation.

Based on the above analysis, this paper proposes the following hypothesis:

H1: The implementation of the Negative List System for Market Access significantly reduces firms' level of real earnings management.

3. Research Design

3.1 Sample Selection and Data Sources

This paper treats the gradual implementation of the Negative List System for Market Access from 2015 to 2019 as an exogenous event and employs a difference-in-differences (DID) model to test the baseline hypothesis. To avoid the impact of the financial crisis on corporate earnings management, this study uses A-share listed firms in the Shanghai and Shenzhen stock exchanges from 2013 to 2019 as the initial sample. The following observations are excluded: (1) ST firms and delisted firms; (2) firms in the financial industry; (3) observations with negative net assets; (4) firms established after 2015; and (5) observations with missing data. In addition, to avoid overstating or distorting the statistical significance of fixed-effects regression results, firms with only a single-period observation are removed. The final sample consists of 15,294 firm-year observations from 2,511 listed firms. The data used in this study are obtained from the China Stock Market and Accounting Research (CSMAR) database.

3.2 Model Specification and Variable Selection

To examine the impact of the multi-period implementation of the Negative List System for Market Access on firms' real earnings management, this paper follows Liang et al. (2021) [12] and Wang (2022) [3] to construct the following fixed-effects model:

$$REM_{i,t} = \beta_0 + \beta_1 Open_{i,t} + \sum Controls + \sum Year + \sum Firm + \varepsilon_{it} \quad (1)$$

The dependent variable, *RealEarningsManage* ($REM_{i,t}$), measures the level of real earnings management. The key independent variable, $Open_{i,t}$, captures whether a firm is affected by the Negative List System for Market Access. If the province where the firm is located is included in the pilot program in a given year, $Open_{i,t}$ equals 1 for that year and subsequent years; otherwise, it equals 0.

The control variables (Controls) include: firm size ($Size_{i,t}$), measured as the natural logarithm of total assets at year-end; leverage ($Lev_{i,t}$), measured as total liabilities divided by total assets at year-end; profitability ($ROE_{i,t}$), measured as net income divided by shareholders' equity; board size ($Board_{i,t}$), measured as the natural logarithm of the number of board members; operating cash flow ($Cfo_{i,t}$), measured as operating cash flow divided by total assets at year-end; proportion of independent directors ($Indep_{i,t}$), measured as the number of independent directors divided by total board size; market-to-book ratio ($MB_{i,t}$), measured as market value divided by book value at year-end; firm age ($Age_{i,t}$), measured as the logarithm of years since listing; earnings management flexibility ($Rev_{i,t}$), measured as the sum of net inventories and net accounts receivable divided by total assets at year-end; and agency cost ($Agency_{i,t}$), measured as administrative expenses divided by operating revenue. The model also controls for firm fixed effects (Firm) and year fixed effects (Year). To mitigate the influence of extreme values, all continuous variables are winsorized at the 1% level on both tails. This paper focuses on the coefficient β_1 of $Open_{i,t}$ in model (1). A significantly positive β_1 indicates that earnings management increases after policy implementation, whereas a significantly negative β_1 indicates that earnings management decreases.

Table 1: Variable Definitions

Variable Symbol	Variable Definition
$REM_{i,t}$	Real earnings management of firm i in year t
$Open_{i,t}$	Dummy variable for the Negative List System for Market Access; equals 1 if the firm's province is under pilot or full implementation in year t , and 0 otherwise
$Size_{i,t}$	Firm size, measured as the natural logarithm of total assets of firm i at the end of year t
$Lev_{i,t}$	Leverage, measured as total liabilities divided by total assets of firm i at the end of year t
$ROE_{i,t}$	Profitability, measured as net income divided by shareholders' equity of firm i in year t
$Board_{i,t}$	Board size, measured as the natural logarithm of the number of board members of firm i in year t
$Cfo_{i,t}$	Operating cash flow, measured as operating cash flow divided by total assets of firm i at the end of year t
$Indep_{i,t}$	Proportion of independent directors, measured as the number of independent directors divided by the total number of board members of firm i in year t
$MB_{i,t}$	Market-to-book ratio, measured as the market value divided by the book value of firm i at the end of year t
$Age_{i,t}$	Firm age, measured as the logarithm of the number of years since firm i was listed
$Rev_{i,t}$	Earnings management flexibility, measured as the sum of net inventories and net accounts receivable divided by total assets of firm i at the end of year t
$Agency_{i,t}$	Agency cost, measured as administrative expenses divided by operating revenue of firm i in year t

4. Empirical Results

4.1 Descriptive Statistics and Univariate Tests

Table 2 reports the descriptive statistics of the main variables. The mean of real earnings management ($REM_{i,t}$) is -0.004 , with a standard deviation of 0.216 , which is close to the results reported by Liang Shangkun. The mean of $Open_{i,t}$ is 0.389 , indicating that 38.9% of the sample experienced the policy shock. In addition, the average leverage ratio is 44.0% , the mean return on assets is 4.6% , the average proportion of independent directors is 37.5% , and the mean logarithm of firm age is 2.4 .

Table 2: Descriptive Statistics

Variable	N	Mean	SD	Min	p25	Median	p75	Max
$REM_{i,t}$	15294	-0.004	0.216	-0.787	-0.099	0.016	0.118	0.626
$Open_{i,t}$	15294	0.389	0.487	0.000	0.000	0.000	1.000	1.000
$Size_{i,t}$	15294	22.393	1.276	19.517	21.508	22.23	23.126	26.135
$Lev_{i,t}$	15294	0.440	0.207	0.049	0.276	0.430	0.594	0.973
$ROE_{i,t}$	15294	0.046	0.161	-1.000	0.024	0.062	0.106	0.436
$Board_{i,t}$	15294	2.136	0.198	1.609	1.946	2.197	2.197	2.708
$Cfo_{i,t}$	15294	0.045	0.068	-0.191	0.008	0.045	0.085	0.241
$Indep_{i,t}$	15294	0.375	0.054	0.333	0.333	0.357	0.429	0.571
$MB_{i,t}$	15294	4.161	3.112	1.251	2.416	3.312	4.716	21.905
$Age_{i,t}$	15294	2.429	0.593	0.959	1.963	2.496	2.970	3.291
$Rev_{i,t}$	15294	0.262	0.165	0.005	0.137	0.242	0.358	0.745
$Agency_{i,t}$	15294	0.095	0.083	0.009	0.045	0.076	0.117	0.617

Table 3 presents the T-test results for differences in means before and after the implementation of the Negative List System for Market Access. The mean of real earnings management ($REM_{i,t}$) is lower after policy implementation and is significant at the 1% level, indicating that real earnings management decreases following the policy. In addition, firm size ($Size_{i,t}$), leverage ($Lev_{i,t}$), board size ($Board_{i,t}$), proportion of independent directors ($Indep_{i,t}$), firm age ($Age_{i,t}$), agency cost ($Agency_{i,t}$), and earnings management flexibility ($Rev_{i,t}$) all exhibit significant differences before and after the policy.

4.2 Hypothesis Testing

Table 4 reports the multivariate regression results of the impact of the Negative List System for Market Access on firms' real earnings management. Column (1) controls only for year and firm fixed effects. The coefficient on the policy variable $Open_{i,t}$ is -0.0154 and statistically significant at the 1% level, indicating that

firms affected by the policy exhibit significantly lower real earnings management compared with unaffected firms. Column (2) further includes firm-level control variables. The coefficient on $Open_{i,t}$ (-0.0124) remains significantly negative at the 1% level, suggesting that the policy effect is robust after controlling for firm size, leverage, profitability, and other characteristics. Column (3) additionally controls for industry–province interaction fixed effects to mitigate potential heterogeneity across regions and industries. The coefficient on $Open_{i,t}$ is -0.0130 and remains highly significant at the 1% level. This indicates that even under a more stringent model specification, the inhibitory effect of the Negative List System for Market Access on real earnings management remains robust.

Table 3: Univariate Analysis

Variable	Before		After		Mean Difference
	Obs	Mean	Obs	Mean	
$REM_{i,t}$	9352	-0.007	1227	-0.025	0.019***
$Size_{i,t}$	9352	22.280	1227	22.446	-0.167***
$Lev_{i,t}$	9352	0.438	1227	0.427	0.012*
$ROE_{i,t}$	9352	0.053	1227	0.066	-0.013***
$Board_{i,t}$	9352	2.145	1227	2.109	0.036***
$Cfo_{i,t}$	9352	0.043	1227	0.041	0.001
$Indep_{i,t}$	9352	0.373	1227	0.380	-0.007***
$MB_{i,t}$	9352	4.525	1227	4.337	0.188*
$Age_{i,t}$	9352	2.355	1227	2.427	-0.072***
$Rev_{i,t}$	9352	0.261	1227	0.273	-0.012**
$Agency_{i,t}$	9352	0.104	1227	0.106	-0.002

Table 4: Baseline Regression Results

	(1)	(2)	(3)
$Open_{i,t}$	-0.0154*** (-2.65)	-0.0124*** (-2.60)	-0.0130*** (-2.70)
$Size_{i,t}$		0.0010 (0.17)	0.0063 (0.89)
$Lev_{i,t}$		0.0431** (2.07)	0.0317 (1.46)
$ROE_{i,t}$		-0.0404*** (-3.59)	-0.0420*** (-3.74)
$Board_{i,t}$		0.0019 (0.12)	-0.0016 (-0.10)
$Cfo_{i,t}$		-1.4735*** (-53.08)	-1.4560*** (-51.35)
$Indep_{i,t}$		-0.0192 (-0.44)	-0.0168 (-0.39)
$MB_{i,t}$		-0.0026*** (-2.98)	-0.0026*** (-2.98)
$Age_{i,t}$		-0.0131 (-1.12)	-0.0134 (-1.09)
$Rev_{i,t}$		0.1497*** (5.67)	0.1986*** (6.73)
$Agency_{i,t}$		-0.0626 (-1.52)	-0.0473 (-1.10)
<i>Cons</i>	0.0021 (0.93)	0.0405 (0.29)	-0.0798 (-0.49)
<i>Year</i>	Yes	Yes	Yes
<i>Firm</i>	Yes	Yes	Yes
<i>Ind#Province</i>	No	No	Yes
<i>N</i>	15294	15294	15269
<i>Adj. R²</i>	0.58	0.72	0.73

*Note: t-values are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. Standard errors are clustered at the firm level. The same applies below.

4.3 Robustness Checks

4.3.1 PSM + DID

To mitigate potential selection bias and sample self-selection issues and to ensure comparability between the treatment and control groups prior to policy implementation, this paper employs the PSM-DID approach for robustness testing. First, the treatment variable (*Treat*) is defined: firms located in provinces that implement the Negative List System for Market Access during the sample period are classified as the treatment group (*Treat*=1), while others serve as the control group (*Treat*=0). Following Liang Shangkun (2021) [12] and related literature on the determinants of earnings management, and based on the research design, the matching variables include firm size (*Size_{i,t}*), leverage (*Lev_{i,t}*), profitability (*ROE_{i,t}*), board size (*Board_{i,t}*), operating cash flow (*Cfo_{i,t}*), proportion of independent directors (*Indep_{i,t}*), market-to-book ratio (*MB_{i,t}*), firm age (*Age_{i,t}*), and earnings management flexibility (*Rev_{i,t}*). Based on these variables, a Logit model is used to estimate propensity scores, and a 1:1 nearest-neighbor matching with a caliper of 0.05 is employed to match treatment firms with similar control firms, ensuring balance in observable characteristics between the two groups.

Using the matched sample, a DID model is then estimated to further identify the net effect of the Negative List System for Market Access on real earnings management. Column (1) of Table 5 reports the PSM-DID regression results. Column (1), which controls for year and firm fixed effects, shows that the coefficient on the core explanatory variable *Open_{i,t}* is -0.0218 and statistically significant at the 1% level. This indicates that after controlling for systematic differences in observable characteristics as well as time and individual heterogeneity, the Negative List System for Market Access still significantly reduces firms' real earnings management, confirming the robustness of the results.

Table 5: PSM-DID and Entropy Balancing Results

	(1)	(2)
<i>Open_{i,t}</i>	-0.0218*** (-2.88)	-0.0155*** (-2.63)
<i>Size_{i,t}</i>	0.0028 (0.35)	-0.0025 (-0.26)
<i>Lev_{i,t}</i>	0.1074*** (4.19)	0.0207 (0.76)
<i>ROE_{i,t}</i>	-0.0384*** (-2.68)	-0.0211 (-1.62)
<i>Board_{i,t}</i>	-0.0106 (-0.51)	-0.0004 (-0.02)
<i>Cfo_{i,t}</i>	-1.4746*** (-37.65)	-1.4681*** (-42.25)
<i>Indep_{i,t}</i>	-0.0471 (-0.75)	-0.0484 (-0.76)
<i>MB_{i,t}</i>	-0.0032** (-2.24)	-0.0015 (-1.28)
<i>Age_{i,t}</i>	-0.0298* (-1.89)	-0.0265 (-1.39)
<i>Rev_{i,t}</i>	0.1026*** (2.64)	0.1783*** (4.96)
<i>Agency_{i,t}</i>	-0.0854* (-1.65)	-0.0515 (-0.94)
Cons	0.0712 (0.39)	0.1769 (0.83)
Year	Yes	Yes
Firm	Yes	Yes
N	7793	15294
Adj. <i>R</i> ²	0.73	0.73

4.3.2 Entropy Balancing Test

To further enhance the credibility of the research findings and effectively mitigate potential selection bias arising from differences in the distribution of covariates between the treatment and control groups, this paper

employs the entropy balancing method for robustness testing. This approach reweights the control group sample to achieve exact balance with the treatment group in terms of the first three moments (mean, variance, and skewness) of the covariates. In doing so, it systematically reduces estimation bias caused by observable differences without imposing linear assumptions on the model specification.

In the empirical implementation, all firm-level characteristics used in the baseline regression are included as covariates, and the policy dummy variable $Open_{i,t}$ is used to define the treatment and control groups to generate balancing weights. These weights are then incorporated into a linear model with both firm and year fixed effects for re-estimation. The regression results are reported in Column (2) of Table 5. The results show that, after entropy balancing, the coefficient on the core explanatory variable $Open_{i,t}$ is -0.0155 and significantly negative at the 1% level. This indicates that even after strictly balancing the distributions of covariates across the treatment and control groups, the Negative List System for Market Access still exerts a significant inhibitory effect on firms' real earnings management.

4.3.3 Parallel Trend Test

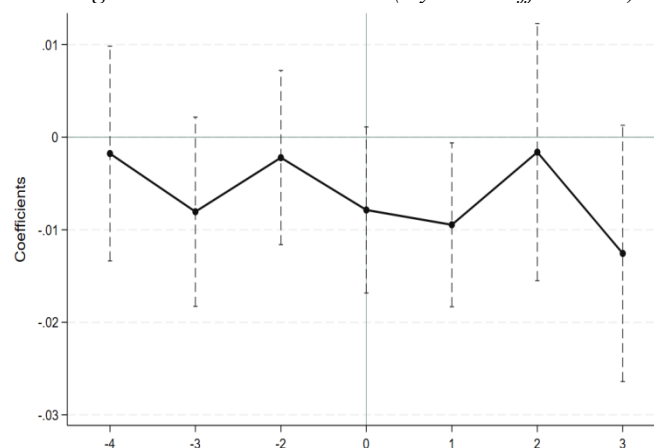
The parallel trend assumption is a key prerequisite for the validity of the difference-in-differences (DID) approach. It requires that, prior to the policy shock, there be no systematic differences in the trends of the outcome variable between the treatment and control groups. To test whether this assumption holds in this study, we follow the existing literature and specify the following event study model:

$$REM_{it} = \alpha + \sum \beta_m Dt-m\ it + \theta X_{it} + \gamma_i + \lambda_t + \varepsilon_{it} \quad (2)$$

where $Dt-m\ it$ denotes a series of event-time dummy variables. The variable m represents the difference between the observation year and the year in which the Negative List System for Market Access is implemented in the province where firm i is located (i.e., the policy shock year). When $m < 0$, it indicates the $|m|$ -th year before the policy shock; when $m = 0$, it denotes the year of the policy shock; and when $m > 0$, it represents the m -th year after the policy shock. The study window is set from four years before to three years after the policy shock, and the period one year prior to the policy shock ($m = -1$) is used as the reference group. The model also controls for firm-level characteristics (X_{it}), firm fixed effects (γ_i), and year fixed effects (λ_t).

Figure 1 reports the estimated coefficients β_m for each event-time dummy variable along with their 90% confidence intervals. The results show that, prior to the policy shock ($m = -4, -3, -2, -1$), all estimated coefficients are close to zero and statistically insignificant, and their confidence intervals include zero. This indicates that there is no significant difference in the trends of real earnings management between the treatment and control groups before the implementation of the Negative List System, thus supporting the parallel trend assumption. Examining the dynamic effects from the year of the policy shock onward, the level of real earnings management in the treatment group shows a downward trend relative to the control group. Specifically, the coefficient for the first period after the policy shock ($m = 1$) is significantly negative at the 10% level, indicating that the policy effect begins to emerge in the short term. Overall, the Negative List System for Market Access exhibits a certain inhibitory effect on firms' real earnings management.

Figure 1: Parallel Trend Test (Dynamic Effects Plot)

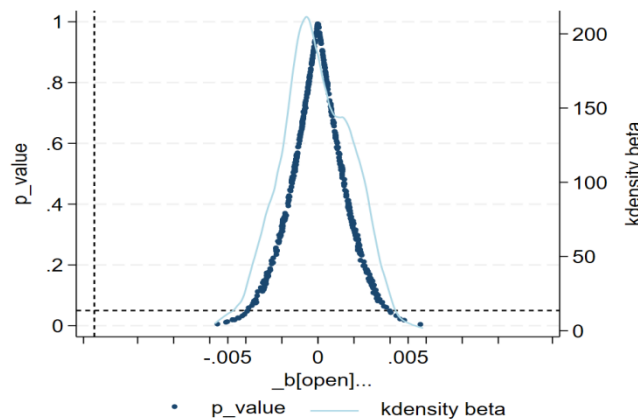


4.3.4 Placebo Test

To rule out the influence of unobservable factors or randomness on the estimation results and to further verify the robustness of the baseline regression findings, this paper follows the existing literature and conducts a placebo test. Specifically, in each year, firms equal in number to the actual treatment group (i.e., firms affected by the Negative List System for Market Access) are randomly selected as a pseudo-treatment group, while the remaining firms serve as a pseudo-control group. A randomly assigned policy dummy variable is then constructed. This process is repeated 500 times, and in each iteration, the same model specification as in the baseline regression is estimated. The estimated coefficients of the policy variable ($Open_{i,t}$) and their significance levels are recorded.

Figure 2 reports the results of the placebo test. Across the 500 random simulations, the distribution of the coefficients on the policy variable ($Open_{i,t}$) is highly concentrated around zero, with a mean approximately equal to 0.000. Compared with the baseline regression coefficient of $Open_{i,t}$ (-0.0114), these estimates are clearly closer to zero. This indicates that the inhibitory effect of the Negative List System for Market Access on real earnings management observed in the baseline regression is not driven by unobservable random factors or omitted variables. Therefore, the placebo test further supports the reliability of the findings in this paper, suggesting that the governance effect of the Negative List System for Market Access on firms' real earnings management is genuine rather than arising from randomness or unobserved confounding factors.

Figure 2: Placebo Test



5. Further Analysis: Heterogeneity Analysis

To further explore the heterogeneous effects of the Negative List System for Market Access on firms' real earnings management, this paper examines two dimensions: corporate internal governance characteristics and financial conditions. Four moderating variables are selected for interaction analysis, including CEO duality ($ConcurrentPosition_{i,t}$), board size ($Board_{i,t}$), operating cash flow ($Cfo_{i,t}$), and profitability ($ROE_{i,t}$). In terms of model specification, each moderating variable and its interaction term with the policy variable ($Open_{i,t}$) are introduced into the baseline regression separately, while controlling for firm, year, and province fixed effects. The regression results are reported in Table 6.

From the perspective of corporate internal governance, CEO duality reflects the degree of managerial power concentration. The results in Column (1) of Table 6 show that the coefficient on the interaction term $ConcurrentPosition_{i,t} \times Open_{i,t}$ is 0.009 and significantly positive at the 10% level, indicating that CEO duality weakens the inhibitory effect of the Negative List System for Market Access on real earnings management. This may be because when the roles of chairman and CEO are held by the same individual, internal checks and balances are weakened, allowing management to better circumvent external institutional constraints, thereby reducing the governance effectiveness of the policy. In contrast, board size reflects the strength of collective oversight. In Column (2) of Table 6, the coefficient on the interaction term $Board_{i,t} \times Open_{i,t}$ is -0.036 and significantly negative at the 1% level, suggesting that the larger the board size, the stronger the inhibitory effect of the policy on real earnings management. A larger board typically implies stronger monitoring capacity and richer decision-making experience, enabling firms to better respond to policy incentives and reduce earnings management behavior.

From the perspective of firms' financial conditions, operating cash flow and profitability capture liquidity pressure and performance pressure, respectively. Column (3) of Table 6 shows that the coefficient on the interaction term $Cfo_{i,t} \times Open_{i,t}$ is 0.012 and significantly positive at the 1% level. As an important indicator of liquidity, firms with abundant cash flows have relatively weaker incentives to engage in real earnings management; therefore, the marginal constraining effect of the policy is smaller, leading to a partial offset of the policy effect. Similarly, in Column (4) of Table 6, the coefficient on the interaction term $ROE_{i,t} \times Open_{i,t}$ is 0.046 and significantly positive at the 5% level. Firms with stronger profitability face less performance pressure and have lower incentives to smooth earnings through real earnings management; thus, the constraining effect of the Negative List System for Market Access is relatively limited for such firms.

Overall, the heterogeneity analysis indicates that the inhibitory effect of the Negative List System for Market Access on firms' real earnings management varies significantly across firms with different governance characteristics and financial conditions. Specifically, the effect is more pronounced in firms with larger boards, while it is relatively weaker in firms with CEO duality, abundant cash flows, and strong profitability. These findings enrich the understanding of the boundary conditions of policy effects and provide useful implications for regulators to implement differentiated policies for different types of firms.

Table 6: Heterogeneity Analysis Results

	(1)	(2)	(3)	(4)
$Open_{i,t}$	-0.016***	0.065**	-0.018***	-0.016**
	(-3.06)	(2.48)	(-3.41)	(-3.08)
$ConcurrentPosition_{i,t} \times Open_{i,t}$	0.009*			
	(1.66)			
$ConcurrentPosition_{i,t}$	-0.009			
	(-1.61)			
$Board_{i,t} \times Open_{i,t}$		-0.036***		
		(-3.00)		
$Board_{i,t}$		0.017		
		(1.02)		
$Cfo_{i,t} \times Open_{i,t}$			0.012***	
			(2.71)	
$Cfo_{i,t}$			-0.009**	
			(-2.34)	
$ROE_{i,t} \times Open_{i,t}$				0.046**
				(2.08)
$ROE_{i,t}$				-0.068***
				(-3.71)
Other Controls	Yes	Yes	Yes	Yes
Cons	0.034	0.018	0.053	0.075
	(0.25)	(0.13)	(0.38)	(0.53)
Yesr	Yes	Yes	Yes	Yes
Province	Yes	Yes	Yes	Yes
N	14861	15294	15294	15294
Adj. R ²	0.725	0.725	0.725	0.725

6. Conclusions, Implications, and Limitations

This paper exploits the phased pilot implementation of the Negative List System for Market Access as a quasi-natural experiment. Using data on A-share listed firms in the Shanghai and Shenzhen stock exchanges from 2013 to 2019 and employing a multi-period difference-in-differences approach, it systematically examines the impact of this system on firms' real earnings management. The results show that the Negative List System for Market Access significantly reduces the level of real earnings management. This conclusion remains robust after a series of robustness checks, including PSM-DID, entropy balancing, parallel trend tests, and placebo tests. Heterogeneity analysis indicates that the policy effect is more pronounced in firms with larger board size, while it is relatively weakened in firms with CEO duality, abundant cash flows, and strong profitability. These findings verify the micro-level governance effects of market-oriented reform from the

perspective of corporate financial behavior, suggesting that lowering entry barriers and strengthening market competition can effectively constrain managerial earnings manipulation.

The above findings yield the following policy implications. First, the reform of the Negative List System for Market Access should be further deepened to fully leverage its role in competition governance. This study confirms that the system can effectively constrain corporate earnings management and improve the quality of financial information. In the future, the dynamic adjustment mechanism of the list should be further improved, and implicit barriers and discriminatory conditions in market entry should be continuously eliminated to create a more fair and transparent competitive environment. Second, greater attention should be paid to firm heterogeneity, and differentiated regulatory strategies should be implemented. Regulatory authorities may focus on firms with smaller boards, CEO duality, abundant cash flows, and strong profitability, as the policy effect is relatively weaker for these firms and complementary regulatory measures are needed. Third, the foundational role of competition policy should be strengthened to give full play to the governance function of market mechanisms. In advancing the reform of “streamlining administration, delegating power, strengthening regulation, and improving services,” fostering a fair competitive environment should be regarded as a key priority, so that market competition can serve as an effective force in constraining opportunistic corporate behavior.

This study has several limitations that warrant further research. First, this paper mainly focuses on identifying the overall policy effect, while the underlying mechanisms require further investigation. Second, due to the limitation of the sample period, this study primarily examines the short-term dynamic effects after policy implementation; future research could employ a longer time horizon to explore the long-term effects. Third, the dimensions of heterogeneity analysis could be further expanded. Factors such as industry characteristics and regional institutional environments may play important moderating roles and deserve more in-depth examination in future studies.

References

- [1] Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Journal of Accounting and Economics*, 42(3), 335–370.
- [2] Cohen, D. A., Dey, A., & Lys, T. Z. (2008). Real and accrual-based earnings management in the pre- and post-Sarbanes–Oxley periods. *The Accounting Review*, 83(3), 757–787.
- [3] Wang, X., & Xu, J. (2022). Does the relaxation of market entry regulation improve firms’ investment efficiency? Evidence from a quasi-natural experiment based on the “negative list” for market access. *Journal of Financial Research*, (9), 169–187.
- [4] Wang, L., Wu, Q., & Luo, D. (2020). Market barriers, administrative approval, and firm price markups. *China Industrial Economics*, (6), 100–117. <https://doi.org/10.19581/j.cnki.ciejournal.2020.06.006>
- [5] Zhang, H., Wang, X., & Zhang, L. (2021). Relaxation of market entry regulation and supply-side capacity reduction: Evidence from a quasi-natural experiment based on the negative list system. *Journal of Finance and Economics*, 47(7), 93–107. <https://doi.org/10.16538/j.cnki.jfe.20210416.403>
- [6] Li, C., Song, M., & Zhang, X. (2014). Analyst coverage and corporate earnings management: Evidence from Chinese listed firms. *Journal of Financial Research*, (7), 124–139.
- [7] Chen, K. (2018). Exit threats and corporate governance: A perspective based on earnings management. *Journal of Finance and Economics*, 44(11), 18–32. <https://doi.org/10.16538/j.cnki.jfe.2018.11.002>
- [8] Li, C., Zhao, Y., Xu, X., & Li, Q. (2016). “Press one spot and another pops up”: Analyst coverage and the choice of earnings management methods. *Journal of Financial Research*, (4), 144–157.
- [9] Xie, D., & Liao, K. (2018). Controlling shareholders’ equity pledges and real earnings management in listed companies. *Accounting Research*, (8), 21–27.
- [10] He, Y., Lin, H., & Wang, M. (2015). Product market competition, executive incentives, and corporate innovation: Evidence from Chinese listed firms. *Finance & Trade Economics*, (2), 125–135. <https://doi.org/10.19795/j.cnki.cn11-1166/f.2015.02.010>

- [11] Chattopadhyay, A., Shaffer, M., & Wang, C. C. (2018). Governance through shame and aspiration: Index creation and corporate behavior. *Journal of Financial Economics* (Forthcoming; European Corporate Governance Institute Working Paper No. 567/2018).
- [12] Liang, S., Cui, H., & Yuan, C. (2021). Stock index constituent adjustments and corporate earnings management: Evidence from a quasi-natural experiment. *Journal of Finance and Economics*, 47(7), 124–138. <https://doi.org/10.16538/j.cnki.jfe.20200616.102>

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

How Management Incentives Influence Corporate Financial Performance: A Mediating Effect Analysis Based on ESG Performance

Jiayi Li*

Accounting and Economics, Beijing Jiaotong University, Beijing, China

**Corresponding author: Jiayi Li.*

Abstract

Against the backdrop of China's dual-carbon targets and the pursuit of high-quality economic development, the synergistic relationship between ESG (environmental, social, and governance) performance and corporate governance mechanisms has become increasingly prominent. Using Chinese A-share listed companies from 2020 to 2024 as the research sample, this paper employs panel-data multiple regression and mediating-effect models to empirically examine both the direct impact of managerial incentives on corporate financial performance and the underlying transmission mechanisms. The results show that (1) both equity-based and compensation-based managerial incentives exert a statistically significant positive effect on corporate financial performance; and (2) ESG performance plays a significant partial mediating role in the relationship between managerial incentives and corporate financial performance—specifically, a transmission pathway exists in which managerial incentives improve ESG performance, which in turn enhances financial performance. These findings remain robust after a series of robustness checks, including variable substitution. By adopting a sustainable-development perspective, the study extends the literature on the economic consequences of managerial incentives. It also supplies empirical evidence that can help firms incorporate ESG metrics into executive evaluation systems, refine incentive contract design, and assist regulators in improving relevant institutional frameworks.

Keywords

ESG performance, managerial incentives, corporate financial performance, mediating effect

1. Introduction

In an era in which China's dual-carbon targets have been fully implemented, the concept of green development has been deeply integrated into market rules, and capital markets are placing increasing emphasis on ESG dimensions in their assessment of corporate sustainability, enterprises must pursue not only short-term financial performance growth to consolidate their market position but also the strengthening of long-term development foundations by improving ESG performance [1].

As a core corporate governance mechanism linking shareholder interests with managerial decision-making, managerial incentives can establish a sustainable pathway for enhancing corporate value by encouraging improvements in ESG performance. This is not only essential for building the competitiveness of individual

enterprises but also holds important demonstration significance for the sustainable development of industries as a whole.

This study selects data from Chinese A-share listed companies from 2020 to 2024 as its research sample. Using panel data regression and mediating effect testing methods, it empirically investigates the impact of managerial incentives on corporate financial performance, with a particular focus on the mediating transmission mechanism of ESG performance. The paper aims to address the shortcomings of existing research on this specific pathway and to enrich the theoretical understanding of corporate governance and long-term value growth. At the practical level, the research findings provide decision-making references for enterprises to optimize incentive contract design and enhance ESG governance levels. They also supply micro-level empirical evidence for regulatory authorities in formulating scientific ESG disclosure and governance policies.

2. Research Hypotheses

2.1 Management Incentives and Corporate Financial Performance

Existing research widely acknowledges the positive governance effects of management incentives on corporate performance. Equity and compensation, as the predominant incentive tools, each operate through distinct mechanisms. As early as 2005, Liu Xia empirically analyzed listed company data, discovering a positive correlation between senior management incentives and corporate operating performance, thereby laying a theoretical foundation for subsequent research into incentive mechanisms [2]. Scholars have also investigated the impact of both equity and compensation, as primary management incentive tools, on corporate financial performance. Jia Zihan et al., focusing on Neusoft Group, confirmed that managerial shareholding in listed companies positively influences corporate financial performance [3]. Yang Xuedan, through empirical research on listed company data, validated a clear positive correlation between managerial compensation and corporate financial performance [4]. These studies consistently demonstrate a positive effect of management incentives on corporate financial performance, and the underlying logic further corroborates this hypothesis. Firstly, equity incentives transform management into de facto "owners" of the company, directly linking their returns to the company's stock price, which helps mitigate short-sighted decisions. Compensation incentives, on the other hand, provide timely feedback on operational results, guiding management to concentrate on core financial indicators such as profitability. Secondly, both equity and compensation incentives attract high-caliber management talent, influencing their interests and risk preferences in innovation activities, thereby promoting the implementation of innovative decisions. This, in turn, enhances corporate financial performance and market value. Based on the aforementioned literature review and logical analysis, the following hypothesis is proposed:

H1: Management equity and compensation incentives empower the enhancement of corporate financial performance.

2.2 Management Incentives, ESG Performance, and Corporate Financial Performance

In recent years, the relationship between management incentives and ESG performance, as well as the link between ESG performance and corporate financial performance, have emerged as prominent research areas. Meng Yun [5] and Liang Wenyu [6] individually investigated the impact of executive equity incentives and high compensation incentives on corporate ESG performance, respectively. Furthermore, research by Wang Xiaoyan et al. affirmed the positive influence of ESG performance on corporate performance [7]. Ran Xu discovered that managerial incentives serve as a mediating pathway in the effect of ESG performance on financial performance [8]. Drawing from these research findings, it can be deduced that management incentives can further influence corporate financial performance by promoting improved ESG performance. Firstly, management incentives can mitigate short-sighted decisions, and since ESG performance constitutes a crucial element of a company's long-term strategy, incentive mechanisms effectively drive enhancements in corporate ESG performance. Secondly, companies with superior ESG performance tend to exhibit better governance capabilities, which is conducive to the development of core financial indicators and the company's long-term progress. Thirdly, increased management incentives encourage companies to prioritize employee rights, thereby fostering greater commitment to social and environmental responsibilities, enhancing ESG performance, and ultimately improving financial performance. In summary, the following hypothesis is proposed:

H2: Management incentives enhance corporate financial performance by promoting improvements in corporate ESG performance.

3. Research Design

3.1 Sample Data Source

This study utilizes annual data from China's A-share listed companies from 2020 to 2024 as the initial sample. ESG performance ratings for A-share listed companies were collected via Tonghuashun. The following operations were then performed: First, samples from the banking and financial sectors were excluded. Second, samples of ST and *ST companies for the respective year were removed. Third, samples with missing data for key variables were excluded to ensure consistency and comparability of financial indicators. This process yielded a total of 25,325 sample observations for parameter estimation.

3.2 Variable Definitions

This study clearly defines and measures the core and control variables involved in the research. Specific details are presented in Table 1.

Table 1: Variable Definition and Measurement

Variable Type	Variable Name	Variable Code	Variable Measurement
Dependent Variable	Corporate Financial Performance	ROE	Return on Equity
Independent Variables	Management Equity Incentive	Share	Ratio of management shareholding to total shares
	Management Compensation Incentive	Salary	Natural logarithm of total management compensation
Mediating Variable	ESG Performance	ESG	Assigned values 1-9 based on Tonghuashun ESG rating (low to high)
Control Variables	Firm Size	Size	Natural logarithm of total assets
	Firm Age	Age	Number of years since firm establishment
	Financial Leverage	Lev	Debt-to-asset ratio
	Growth	Gro	Operating revenue growth rate
	Equity Concentration	Top5	Sum of shareholding percentages of the top five shareholders
	Number of Managers	Num	Total number of managers

3.3 Model Construction

Based on the preceding analysis, the basic model is constructed as follows:

$$Y = \alpha_0 + \alpha_1 X + \sum \alpha_j C + \varepsilon_1 \quad (1)$$

$$M = \beta_0 + \beta_1 X + \sum \beta_j C + \varepsilon_2 \quad (2)$$

$$Y = \gamma_0 + \gamma_1 X + \gamma_2 M + \sum \gamma_j C + \varepsilon_3 \quad (3)$$

In the formula, α , β , and γ represent parameters, X denotes the dependent variable (corporate financial performance), Y represents the independent variable (management incentives), M denotes the mediating variable (ESG performance), and C represents the control variables.

4. Empirical Results and Analysis

4.1 Descriptive Analysis of Variables

The descriptive statistics for the main variables are presented in Table 2. The mean value of Corporate Financial Performance (ROE) is 0.035, with a standard deviation of 0.098, indicating significant variation in the financial performance (ROE) among the sample A-share listed companies. A comparison of the mean and

standard deviation for Management Equity Incentive (Share) and Management Compensation Incentive (Salary) reveals substantial differences in both equity and compensation incentive levels. The mean ESG performance is 5.193, suggesting that the average ESG rating of the sample companies falls within the BB-BBB range. The descriptive statistics for the remaining control variables are shown in the table below:

Table 2: Descriptive Statistics of Main Variables

Name	Sample Size	Minimum Value	Maximum Value	Mean	Standard Deviation	Median
Roe	25325	-2.49	7.44	0.03	0.098	0.03
Share	25325	0.00	0.90	0.17	0.207	0.05
Salary	22858	11.77	18.93	15.72	0.657	15.69
ESG	22642	1.00	9.00	5.19	1.362	5.00
Age	25325	2.00	71.00	21.02	6.747	21.00
Size	25325	17.87	28.79	22.20	1.403	21.97
Lev	25325	0.01	3.64	0.39	0.208	0.38
Top5	25325	2.69	98.63	54.26	16.319	53.92
Gro	25325	-1.44	944.10	0.30	9.532	0.07
Num	25325	8.00	53.00	17.93	4.884	17.00

4.2 Analysis of Baseline Regression Results

A linear regression model was employed for the analysis, and the results are presented in the tables below. Table 3 shows that the regression coefficient of Share on ROE is 0.053 ($t=17.663$, $p=0.000 < 0.01$), indicating that Share (X1) has a significant positive impact on ROE (Y). Table 4 demonstrates that the regression coefficient of Salary on ROE is 0.016 ($t=15.817$, $p=0.000 < 0.01$), meaning that Salary (X2) also has a significant positive impact on ROE (Y). The overall regression results align with expectations, thus, the baseline regression results validate Hypothesis H1.

Table 3: Linear Regression Results for Share

	Non-Standardized Coefficients	Standardized Coefficients	t	p	Collinearity Diagnostics
	B	Std. Error	Beta		VIF
(Constant)	0.027	0.001	-	33.276	0.000**
Share (X1)	0.053	0.003	0.110	17.663	0.000**
R ²	0.012				
Adjusted R ²	0.012				
F	F (1,25323) = 311.991, $p=0.000$				
D-W Value	1.486				

Note: Dependent Variable = ROE (Y), * $p < 0.05$ ** $p < 0.01$

Table 4: Linear Regression Results for Salary

	Non-Standardized Coefficients	Standardized Coefficients	t	p	Collinearity Diagnostics
	B	Std. Error	Beta		VIF
(Constant)	-0.220	0.016	-	-13.853	0.000**
Salary (X2)	0.016	0.001	0.104	15.817	0.000**
R ²	0.011				
Adjusted R ²	0.011				
F	F (1,22856) = 250.188, $p=0.000$				
D-W Value	1.521				

Note: Dependent Variable = ROE (Y), * $p < 0.05$ ** $p < 0.01$

4.3 Mediating Effect Analysis

Table 5 presents the results of the mediating effect test of ESG in the relationship between Share and Salary on ROE. Column (1) indicates that both Share and Salary have a significant positive impact on ROE. Column (2) demonstrates that both Share and Salary also have a significant positive impact on ESG. Column (3) shows that while Share and Salary have a positive impact on ROE, ESG also has a significant positive impact on ROE. Furthermore, the significance of the positive impact of Share and Salary on ROE remains largely

unchanged, still showing strong significance. Therefore, ESG plays a partial mediating role in the influence of Share and Salary on ROE, supporting Hypothesis H2, and the results are relatively robust.

Table 5: Mediating Effect Analysis Results

	(1)	(2)	(3)
Variables	Roe	ESG	Roe
Share	0.023** (0.003)	0.277** (0.045)	0.021** (0.003)
Salary	0.014** (0.001)	0.381** (0.015)	0.011** (0.001)
ESG			0.007** (0.001)
Controls	Yes	Yes	Yes
R ²	0.145 (0.149)	0.224 (0.244)	0.152 (0.155)
N	22559	22559	22559

Note: * $p < 0.05$ ** $p < 0.01$

4.4 Robustness Analysis

To further verify the robustness of the baseline regression and mediation effect results, and to mitigate potential endogeneity concerns arising from measurement bias in single indicators, this study conducts robustness tests using a variable substitution approach across three dimensions: the explanatory variable, the dependent variable, and the mediating variable.

1. Substitution of the explanatory variable.

In empirical analysis, the measurement of incentive variables may be sensitive and thus affect estimation results. Accordingly, this study replaces the original explanatory variable with two alternative measures: (i) the *frequency of managerial equity incentives* and (ii) the *natural logarithm of the total compensation of the top three executives*, and re-estimates the regressions. The frequency of equity incentives effectively captures the continuity and intensity of long-term interest alignment mechanisms, while the compensation of the top three executives better reflects the incentive strength within the firm's core decision-making group. The results show that, after these substitutions, managerial incentives continue to exert a statistically significant and positive effect on firm financial performance. This indicates that the performance-enhancing effect of incentive contracts does not depend on a specific measurement, thereby confirming the robustness of Hypothesis H1.

2. Substitution of the dependent variable.

Compared with traditional accounting-based indicators such as ROE, which primarily reflect historical or contemporaneous book performance, Tobin's Q incorporates the capital market's forward-looking assessment of a firm's growth potential and intangible assets. When financial performance is proxied by Tobin's Q, the direction and statistical significance of the core variables remain unchanged. This not only confirms that managerial incentives improve short-term accounting performance, but also suggests that, by enhancing ESG performance, firms convey positive signals of sustainable development to external markets, thereby obtaining a higher market valuation premium.

3. Substitution of the mediating variable.

Considering that the ESG rating used in the baseline regression (a 1–9 scale) is a categorical variable, and to avoid potential bias in identifying the mediation effect due to measurement precision, this study replaces the mediating variable with a more granular continuous measure—the *Tonghuashun ESG composite score*—and re-examines the model. The results indicate that, with improved measurement precision, ESG performance continues to play a significant partial mediating role between managerial incentives and financial performance, further reinforcing the validity of the transmission mechanism of “incentive-driven → ESG improvement → performance enhancement.”

Overall, after systematically implementing multidimensional variable substitutions, the signs and significance levels of the key coefficients remain largely unchanged compared to the baseline model. This

provides strong evidence that the findings of this study are robust and generalizable, and are not driven by arbitrary choices of measurement.

5. Conclusion

This study utilizes data from A-share listed companies in China from 2020 to 2024 to empirically examine the impact of management equity incentives and compensation incentives on corporate financial performance and their underlying mechanisms. The research reveals that both incentive methods significantly enhance financial performance, with ESG performance playing a crucial partial mediating role. This confirms the effectiveness of the "incentive-driven—ESG optimization—performance enhancement" transmission chain. Further comparison indicates that equity incentives (Share) have a more prominent direct boosting effect on financial performance. This is primarily attributed to the long-term lock-in effect of equity, which inclines management to maintain corporate valuation from a strategic perspective. Conversely, compensation incentives (Salary) demonstrate a more significant marginal promoting effect on ESG performance. This reflects that, under the current appraisal system, management is more sensitive to cash incentives tied to annual performance and is more readily motivated to improve ESG practices in the short term to gain immediate feedback. This differentiated performance suggests that equity incentives are more focused on results-oriented value creation, while compensation incentives are more efficient in driving management to fulfill process-oriented behaviors, such as social responsibility.

Based on these findings, enterprises should optimize incentive contract design by deeply embedding ESG indicators into their appraisal systems. By combining "equity for long-term effectiveness and compensation for behavioral promotion," companies can strengthen the empowering role of incentive mechanisms for sustainable development. Concurrently, regulatory authorities should further improve information disclosure systems to enhance incentive transparency and strengthen market supervision. Although this paper provides valuable empirical evidence, certain limitations remain. Firstly, constrained by data availability, this study primarily focuses on A-share listed companies, and its generalizability to unlisted or small and medium-sized enterprises requires further validation. Secondly, ESG performance is heavily influenced by macro policies and industry attributes. While this paper controlled for factors such as time, it could not exhaust all external environmental variables. Future research could attempt to introduce an industry-segmentation perspective or integrate methods such as machine learning to conduct deeper investigations into the non-linear impacts of ESG performance.

References

- [1] Sun, N., & Qu, W. H. (2024). ESG performance empowers new quality productive forces of enterprises under the "double carbon" goal. *Statistics & Information Forum*, 39(10), 24–41. (in Chinese)
- [2] Liu, X. (2018). An empirical study on the correlation between senior management incentives and the operating performance of listed companies. *Modern Marketing*, (30), 195. (in Chinese)
- [3] Jia, Z. H., Zhu, J. Z., & Wang, W. T. (2025). A case study on the impact of equity incentives on the financial performance of listed companies: Based on the dual perspective of incentive modes and exercise conditions. *Friends of Accounting*, (S2), 136–143. (in Chinese)
- [4] Yang, X. D. (2021). Empirical research on the relationship between executive compensation incentives and corporate financial performance: Based on the perspective of corporate governance. *Market Modernization*, (11), 169–171. (in Chinese)
- [5] Meng, Y. (2024). *Research on the impact of executive equity incentives on corporate ESG performance* [Master's thesis/Doctoral dissertation, Liaoning University]. (in Chinese)
- [6] Liang, W. Y. (2024). *Research on the impact of executive compensation incentives on corporate ESG performance* [Master's thesis/Doctoral dissertation, Shanxi University of Finance and Economics]. (in Chinese)

- [7] Wang, X. Y., & Zhou, D. Q. (2026). The impact of ESG performance on corporate financial performance: An empirical test based on the perspective of China's manufacturing industry. *South China Finance*, 1–17. (in Chinese)
- [8] Ran, X. (2024). ESG performance, management incentives, and corporate financial performance. *Old-Established Brand Marketing*, (22), 127–129. (in Chinese)

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Money Intuition in Digital Financial Behavior: The Impact Mechanism of Mobile Payment and Credit Consumption on Financial Risk Perception among Quasi-Digital Natives

Jiyu Yuan*

Business School, Nankai University, Tianjin 300071, China

**Corresponding author: Jiyu Yuan.*

Abstract

The deep popularization of digital payment has profoundly influenced the consumption behavior and money cognition of young people, but it has also raised concerns about the weakening of financial risk perception. This paper introduces the core concept of “money perceptibility,” which refers to individuals’ subjective judgment of the authenticity of money’s value and their comprehensive psychological perception during spending. Based on this, the paper constructs a theoretical analysis framework of “payment method—money perceptibility—financial risk perception.” It explores the impact mechanisms of mobile payment and credit consumption on the financial risk perception of quasi-digital natives through three pathways: sensory, cognitive, and emotional. Accordingly, the paper proposes countermeasures from three levels—family and school education, corporate product design, and government regulatory protection—aiming to help young people reshape healthy money cognition and enhance their ability to guard against financial risks in the digital era.

Keywords

digital payment, money perceptibility, financial risk perception, quasi-digital natives, payment pain, credit consumption

1. Introduction

1.1 Statement of the Problem

The rapid development of the digital economy has become a key feature of China’s economic and social transformation (the opening of the “14th Five-Year Plan for Digital Economy Development” states that “the digital economy is the main economic form following the agricultural and industrial economies”). The high-speed development of the information age has also led to the rapid growth of digital payment, which now dominates the consumption behavior of people in daily life, especially among young people. According to data from the People’s Bank of China, banks processed 210.98 billion mobile payment transactions in a year, with a transaction value of 563.70 trillion yuan. The payment system operated smoothly, and the volume of non-cash payment business continued to increase steadily [1,2]. Additionally, the 54th “Statistical Report on

Internet Development in China” released by the China Internet Network Information Center (CNNIC) shows that as of June 2024, China had nearly 1.1 billion netizens (1.09967 billion), of whom 969 million used online payment services, accounting for 88.1% of all netizens [3]. This indicates that digital payment has deeply integrated into economic and social life and has become a core force driving the development of the digital economy. The user penetration rate of credit consumption tools such as Pinduoduo’s “pay later” and Alipay’s Huabei continues to rise. The digital transformation of securities trading and online lending has further simplified large-sum fund transfers into mere changes in numbers on a screen. The physical attributes of money are gradually being dissolved, and digital symbols have become the primary form through which contemporary young people perceive money. At the same time, this raises questions about contemporary young people’s perception of financial risk. Does this digitized way of understanding money weaken young people’s risk sensitivity?

Although existing studies have confirmed the impact of payment methods on consumption behavior and the influence of credit consumption on debt risk, and have proposed related theories such as “payment pain” and “mental accounting,” significant research gaps remain. Most existing studies focus on post-90s and post-00s groups, with relatively little attention paid to the “quasi-digital natives”—the late Z generation and early Alpha generation (born between 2005 and 2015)—who grew up highly dependent on digital payment and had limited exposure to physical currency [4,5]. Based on this, the present paper takes the “quasi-digital natives” as the research object, proposes the core analytical concept of “money perceptibility,” systematically analyzes the internal mechanisms by which digital payment and credit consumption affect young people’s financial risk perception, identifies potential risks, and offers targeted recommendations. It seeks to answer the core question: “How does the digital era shape young people’s money intuition?”

1.2 Research Significance

1.2.1 Theoretical Significance

The theoretical contributions of this paper can be summarized in three aspects.

First, this study is the first to systematically and rigorously define the concept of “money perceptibility” as individuals’ comprehensive perceptual ability regarding the physical attributes, value attributes, and risk attributes of money. This effectively fills the theoretical gap in existing literature concerning the money cognition of digital natives and opens a new perspective for research on the relationship between digital payment and financial risk perception [6].

Second, by integrating the “payment pain” theory from behavioral economics, the embodied cognition perspective from psychology, and sociological analyses of the generational characteristics of “digital natives,” this paper constructs a theoretical framework suitable for China’s highly digitalized context, thereby providing a more solid understanding of “how digital payment reconstructs individuals’ money cognition patterns” [4,6].

Third, by conducting a clear and insightful examination of the “quasi-digital natives” group, this paper breaks through the long-standing limitation in relevant research that has largely overlooked this cohort and focused primarily on post-90s and post-00s adult users. It not only expands new empirical samples for consumption behavior research from an intergenerational difference perspective but also lays a foundation for future studies on the financial behavior of late Z generation and early Alpha generation individuals.

1.2.2 Practical Significance

From a practical perspective, this study offers three layers of real-world value.

First, by clarifying the various consequences of the lack of “money perceptibility,” the paper provides references for financial institutions, educational platforms, and regulatory authorities in designing financial literacy courses, risk warning tools, and related policies, thereby effectively helping “quasi-digital natives” develop healthy money intuition and rational consumption habits.

Second, in optimizing the design of digital financial products and risk control mechanisms, it must be recognized that mainstream payment and credit products such as Huabei, JD White Bar, and “pay later” services deliberately weaken users’ sense of payment pain during design, thereby diminishing users’ perception of their own assets and inadvertently inducing debt tendencies. The concept of “money

perceptibility” proposed in this paper can naturally serve as a new indicator for evaluating product experience, guiding enterprises to appropriately retain “payment salience” while ensuring service convenience, and truly achieving a balance between commercial benefits and user rights.

Third, from the perspective of national digital economy governance and financial stability strategy, as the penetration rate of digital payment continues to increase, issues such as debt risk and irrational investment among young people have become more prominent. Starting from micro-level psychological mechanisms, this paper provides clear and insightful analysis on “how the digital environment shapes the consumption behavior of the new generation,” thereby offering policy recommendations for preventing systemic financial risks.

2. Literature Review and Conceptual Framework

2.1 Research Status and Theoretical Gaps

The impact of payment methods on individual consumption behavior is a classic and well-researched topic in behavioral economics and consumer psychology. Existing literature is primarily based on two major theoretical frameworks: the “Pain of Paying” theory and the “Mental Accounting” theory. The “Pain of Paying” theory, proposed by Prelec and Loewenstein (1998), presents a clear and powerful logical chain: the physical attributes of the payment medium directly affect consumers’ perception of payment pain. Cash payments require physical handover and bill counting, allowing consumers to directly and fully perceive the reduction in their assets. This process activates brain regions associated with pain and thereby inhibits impulsive consumption [4,5]. In sharp contrast, non-cash payment methods such as credit cards and digital payments, due to their abstract, non-immediate, and lack of physical contact nature, significantly weaken this pain sensation, leading to increased spending. More recent studies have further rigorously verified that mobile payment reduces payment salience more than traditional card swiping, causing consumers to more easily overestimate their purchasing power and underestimate total expenditure.

Literature in the field of credit consumption has appropriately applied Thaler’s “Mental Accounting” theory to explain young people’s debt behavior. Tools such as Huabei and JD White Bar, which operate on a “buy now, pay later” model, cleverly exploit the separation between the “consumption moment” and the “payment moment,” reconstructing users’ mental accounts. At the moment of consumption, users only perceive the use of credit limits rather than the depletion of their own assets, deferring repayment pressure to a future point. This implicit feature makes it easier for users to underestimate the actual cost of debt [5]. Complementing this, intergenerational difference research on “Digital Natives” notes that this group grew up in an internet environment with a natural affinity for digital technology. However, current academic research has mostly focused on post-90s and early post-00s groups, with relatively little attention paid to the “quasi-digital natives” born between 2005 and 2015 [6]. More importantly, this group is in a critical transitional period from “having cash memory” to “having no cash memory,” so the logic of their money cognition formation differs fundamentally from that of previous generations.

Domestic scholars have conducted solid and multi-layered localized research on this topic. Liu Wenchang (2026) argued that digital payment reduces payment pain and reconstructs mental accounts, thereby leading to fragmented and irrational consumption decisions [5,7]. Nevertheless, existing literature still has two clear and important research gaps: first, the research subjects are incomplete, with no dedicated analysis of “quasi-digital natives”; second, mechanism exploration is not sufficiently deep. Current studies mostly discuss consumption amounts directly from payment methods without clarifying the psychological transmission mechanisms—namely, individuals’ psychological perception process of the authenticity of money—which is precisely the research value of the core concept “money perceptibility” defined in this paper.

2.2 Definition of the Core Concept: Money Perceptibility

Addressing the research deficiencies mentioned above, this paper proposes “Money Perceptibility” as the core analytical concept. It does not refer to the physical form of currency itself, but rather to individuals’ comprehensive subjective experience of the authenticity, scarcity, and psychological pain associated with money’s value during spending situations. In short, it is the intensity of individuals’ psychological perception of the authenticity of money as a resource.

Specifically, this paper divides “money perceptibility” into two typical states:

(1) High Money Perceptibility:

Individuals have a clear perception of the amount and value of their held funds and regard them as limited resources. In this state, individuals experience strong psychological pain during spending, and their consumption decisions tend to be cautious and rational. This typically corresponds to scenarios with strong physical feedback, such as cash payments.

(2) Low Money Perceptibility:

Individuals have a vague perception of the amount and value of their held funds and tend to view money as digital symbols without clear boundaries. In this state, individuals experience very low or even no psychological pain during spending, easily overlook the scarcity and potential risks of money, and are prone to impulsive consumption decisions. This typically corresponds to scenarios with weak physical feedback, such as mobile payment and credit payment.

This paper argues that digital payment reduces young people’s “money perceptibility” by stripping away the physical attributes of money, thereby leading to a blunting of their financial risk perception. This concept integrates the emotional dimension of “payment pain” from behavioral economics and the cognitive dimension of “money form effects,” providing a new theoretical perspective for explaining the irrational tendencies in young people’s financial behavior in the digital era [4,8].

2.3 Theoretical Model and Research Hypotheses

Based on the definition and deconstruction of the core concept of “money perceptibility,” this paper posits that digital payment and credit consumption are not direct causes of young people’s risky financial behavior. Instead, they exert their effects through a hidden psychological mediating mechanism—the weakening of “money perceptibility.” This process constitutes the core analytical logic of this paper and can be broken down into three deductive links:

First, the evolution of payment methods reshapes the foundation of individuals’ money perception. Traditional cash payments possess distinct physical attributes. The tactile sensation of paper money, the weight of coins, and the actions of counting create a multi-sensory “embodied cognition” experience that allows individuals to intuitively confirm the reduction of assets and maintain high money perceptibility. However, digital payment abstracts money into symbols on a screen, severing the multi-sensory feedback during the payment process. Credit payment tools such as Huabei and JD White Bar further separate “consumption” from “payment” in time, blurring the boundaries of mental accounts [5]. For “quasi-digital natives,” who have grown up almost entirely in a cashless environment, the de-physicalized and de-immediate characteristics of these payment methods reduce their perception of the authenticity and scarcity of money, significantly lowering “money perceptibility.”

Second, the absence of money perceptibility is the key trigger for the blunting of financial risk perception. When individuals are in a state of “high money perceptibility,” every expenditure is accompanied by a clear “sense of loss” and psychological resistance. This instinctive emotional response forms the first line of defense against financial risk, prompting cautious decision-making. Conversely, in a state of “low money perceptibility,” spending becomes “painless” and “imperceptible.” In this psychological state, individuals’ vigilance toward capital outflows and debt accumulation decreases significantly. In other words, it is precisely the lack of “money perceptibility” that makes it difficult for young people to generate appropriate risk anxiety when facing high-value consumption or credit temptations, leading to a decline in risk perception.

Third, generational characteristics amplify this transmission mechanism. Compared with post-90s or earlier generations, quasi-digital natives have a highly distinctive growth trajectory: they are in a critical transitional period from “having cash memory” to “having no cash memory,” and some have never even developed experience with physical currency. The absence of this “experience anchoring” with cash makes them more adaptable to the abstract features of digital payment, but it also means their money perceptibility is more easily weakened by the digital environment. Therefore, in the transmission chain of “payment

method → money perceptibility → risk perception,” the effect observed among “quasi-digital natives” will be more significant and thorough than in previous generations.

3. Mechanism Analysis: Pathways Through Which Digital Payment and Credit Consumption Affect Money Perceptibility and Financial Risk Perception

Digital payment and credit consumption do not directly affect individuals’ rational decision-making capacity. Instead, they systematically and hierarchically erode the “money perceptibility” of quasi-digital natives through three interconnected psychological pathways—sensory, cognitive, and emotional—thereby blunting financial risk perception. This outcome is essentially the result of the interaction of multiple mechanisms.

3.1 Sensory Pathway: Discussion on the Stripping of Money’s Physical Attributes and the Breakage of the Embodied Cognition Chain

In traditional cash transactions, money possesses distinct and perceptible physical attributes: the texture of paper bills, the weight of coins, the sound of counting, and the hand movements during delivery. These factors together constitute a complete sensory experience of cash transactions. Visual, tactile, and auditory stimuli interconnect to form individuals’ typical “embodied cognition” of the existence of money. More importantly, behavioral neuroscience has provided clear and rigorous evidence that physical contact activates brain regions in the insular cortex directly associated with pain and loss, thereby generating an intuitive “pain of paying” and forming a physiological barrier that inhibits impulsive consumption [4].

However, the emergence of digital payment completely severs the original sensory feedback chain: payment is simplified to changes in numbers on a screen, fingerprint, or facial recognition. Consequently, the physical carrier of money has been replaced by abstract codes. For “quasi-digital natives,” the abstraction of digital payment is not merely a change in payment method; it is fundamentally a systematic reconstruction of the underlying logic of money cognition. Since they have had little exposure to the physical feedback of cash decreasing transaction by transaction, account balances in their cognition resemble a set of digital symbols without clear boundaries. Therefore, it is reasonable to introduce the concept of “de-physicalization”: it directly leads to a substantial increase in the “payment pain” threshold [4]. Without immediate sensory feedback, it becomes difficult to generate psychological resistance during spending, making capital outflows painless and imperceptible. In other words, sensory numbness is the primary pathway for the weakening of money perceptibility and serves as the physiological foundation for the blunting of financial risk perception.

3.2 Cognitive Pathway: Blurring of Mental Accounts and Distortion of Value Perception

Beyond the stripping at the sensory level, digital payment and credit consumption also actively and systematically blur the boundaries of “mental accounts,” thereby exerting a clear influence on individuals’ cognition of money’s value [5]. Thaler’s mental accounting theory illustrates this point: people spontaneously categorize funds from different sources and for different purposes into separate mental accounts and apply different spending rules to each. In the cash transaction era, the sources of funds were clear (salary, pocket money) and their purposes well-defined (salary for savings, pocket money for entertainment), resulting in distinct account boundaries. Expenditure directly meant an absolute reduction in the balance of a particular account [5].

In digital credit consumption scenarios, however, the original stable cognitive structure is broken. Tools such as Huabei, JD White Bar, and “pay later” services clearly separate “current consumption” from “future repayment” in both time and space. Thus, at the moment of consumption, users perceive only the use of credit limits rather than the depletion of their own assets, easily creating the illusion of “not spending their own money” and readily categorizing the expenditure into a “future account.” This leads to the typical phenomenon of blurred mental account boundaries characterized by the “separation of consumption from payment,” with the cost of funds being rendered implicit and deferred [5]. For “quasi-digital natives,” whose cognitive control abilities are still developing, this ambiguity easily causes distortion in value perception: underestimating the true cost of current consumption, overestimating future repayment capacity, and transforming money from a scarce resource into an easily accessible digital permission. More seriously, this cognitive bias weakens money perceptibility, making it extremely difficult for individuals to establish a

direct and clear logical connection between consumption and debt when facing large expenditures or liabilities, and thus extremely difficult to arouse corresponding risk vigilance.

3.3 Emotional Pathway: Reinforcement of Instant Gratification and Lag of Anxiety

The combination of digital payment and credit consumption creates an emotional feedback loop of “low resistance, high temptation.” Features such as one-click ordering, interest-free installments, and instant fund arrival greatly shorten the time gap between “desire arousal” and “desire satisfaction,” thereby facilitating greater dopamine release during instant gratification. Consequently, consumption behavior gradually becomes a direct means for people to regulate emotions and obtain pleasure under high-frequency positive emotional reinforcement, rather than merely a resource allocation activity.

Negative emotions such as anxiety and worry triggered by financial risk exhibit clear and identifiable lag: due to repayment cycles, debt pressure does not appear immediately upon consumption but is systematically postponed to a future point. This is particularly typical for young people who lack long-term planning experience. Over time, individuals’ emotional sensitivity to financial risk gradually dulls, forming a “blindly optimistic” risk preference. More importantly, due to the lack of “money perceptibility,” individuals are actually unable to truly “feel” the existence of risk at the emotional level. The three pathways are interconnected and mutually causal, thus revealing the deep psychological mechanisms underlying the blunting of financial risk perception among young people in the digital era.

4. Conclusion and Recommendations

4.1 Main Conclusions

Starting from the macro background of the rapid development of the digital economy, this paper takes quasi-digital natives as the research object and “money perceptibility” as the basic concept. It constructs the theoretical framework of “payment method—money perceptibility—financial risk perception” and systematically and rigorously examines the impact mechanisms of digital payment and credit consumption on financial risk perception [6].

Money perceptibility is a clear and important psychological variable explaining the alienation of young people’s financial behavior in the digital era. This paper demonstrates a positive correlation between the tangibility of payment methods and individuals’ money perceptibility: digital payment and credit payment strip away the physical attributes of money, simplify the payment process, and transform money into abstract symbols, thereby directly weakening individuals’ concrete perception of money’s value.

From the existing arguments, it can be clearly and rigorously observed that money perceptibility plays a full mediating role between payment methods and financial risk perception. Digital payment and credit consumption do not directly cause a decline in risk perception; rather, they first reduce money perceptibility, which naturally and reasonably blunts individuals’ vigilance toward financial risk. For “quasi-digital natives,” due to the lack of experience anchoring with cash, this mediating effect is particularly pronounced, making them more likely to make irrational consumption decisions during credit consumption.

4.2 Policy Recommendations

Based on the above conclusions, this paper proposes targeted recommendations from three levels: micro, meso, and macro.

First, at the family and school level:

Financial literacy education should be incorporated into compulsory education curricula. Adolescents should be encouraged to keep accounts and visualize expenditures to deliberately train the correspondence between digital amounts and real purchasing power. Parents can appropriately retain part of pocket money in cash form, allowing students to manage cash personally, artificially reconstructing “payment pain” [4], cultivating healthy money intuition, and enhancing the ability to guard against financial risks.

Second, at the corporate level: Adhere to product design ethics.

Enterprises should balance commercial interests with user rights. For example, multiple confirmations should be required for large-value consumption; a “cooling-off period” function should be set for users whose debt exceeds a certain limit, displaying the current repayment amount required and allowing payment only after 24 hours, during which the order can be canceled at any time. This moderately awakens users’ risk awareness, avoids algorithm-induced excessive borrowing, and achieves a balance between commercial benefits and user well-being.

Third, at the regulatory level: Strengthen protection for adolescent finance.

Regulatory authorities should, based on the new situation of digital financial development, establish stricter credit consumption access and quota management systems for quasi-digital natives and severely crack down on inductive marketing targeted at minors. Efforts should be made to promote a healthy digital financial ecosystem, prevent systemic financial risks from the source, and promote the healthy and sustainable development of the digital economy.

4.3 Research Limitations and Future Prospects

Although this study has made certain progress in theoretical construction and mechanism analysis, it still has limitations. The research methods mainly consist of theoretical deduction and mechanism analysis, lacking empirical data support. Future research can design targeted questionnaires for surveys to further verify the rationality of this study, and conduct laboratory experiments to compare individuals’ money perceptibility and financial risk perception under different payment methods. This will explore specific pathways to enhance young people’s financial risk perception and provide more solid empirical support for policy formulation.

References

- [1] State Council of the People’s Republic of China. (2021). *The 14th Five-Year Plan for Digital Economy Development* [Policy document]. Beijing: State Council.
- [2] People's Bank of China. (2025). *Overall operation of the payment system in 2024* [Report]. Beijing: People's Bank of China.
- [3] China Internet Network Information Center. (2024). *The 54th Statistical Report on Internet Development in China* [Report]. Beijing: China Internet Network Information Center.
- [4] Prelec, D., & Loewenstein, G. (1998). The Red and the Black: Mental Accounting of Savings and Debt. *Marketing Science*, 17(1), 4–28.
- [5] Richard H. Thaler. (2008). Mental Accounting and Consumer Choice. *Marketing Science*, 27 (1), 15-25.
- [6] Marc Prensky. (2001). Digital Natives, Digital Immigrants Part 1. *On the Horizon*, 9(5), 1-6. <https://doi.org/10.1108/10748120110424816>.
- [7] Liu, W. C. (2026). Research on the impact of digital payment on residents' consumption level and structure in China. *Shanghai Commerce*, (02), 10–11.
- [8] Raghurir, P., & Srivastava, J. (2009). The denomination effect. *Journal of Consumer Research*, 36(4), 701-713.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Digital Twin Technology Drives Real-Time Scheduling Optimization for Port Container Operations: A Verification Study Centred on the Goals of Efficiency Improvement and Cost Control

Minghan Huang*

Supply Chain Management Department, Northeast Normal University, Changchun 130024, China

**Corresponding author: Minghan Huang.*

Abstract

In the context of global port intelligent transformation, digital twin technology has become a key support for breaking through the bottleneck of container operation scheduling efficiency and reducing operational costs. This study adopts a single-case exploratory design, selecting typical ports that have deployed digital twin systems as the research objects. It systematically examines the application mechanism of digital twin technology in container operation scheduling from three dimensions: technical empowerment, decision reconstruction, and performance optimization. The research team obtained multiple data sources through field investigation methods such as participatory observation, system operation logs, scheduling record reports, and in-depth interviews. They used the three-level coding method of grounded theory to construct a process model of “technical empowerment - decision transformation - performance output”, revealing the complete transformation path from technical input to efficiency-cost dual-target output. The study found that the digital twin system realizes real-time integration and virtual mapping of all process data through a hierarchical architecture, forming a hierarchical human-machine collaborative decision-making model of “automated in routine scenarios, collaborative in complex scenarios, and manual in emergency scenarios”, effectively achieving the collaborative optimization of vessel operation efficiency and reduction of operational costs. The process model constructed in this study reveals the internal mechanism of digital twin technology driving scheduling optimization, providing systematic theoretical and practical guidance for the construction of intelligent port scheduling.

Keywords

digital twin technology, port container operation scheduling, human-machine collaborative decision-making

1. Introduction

With the acceleration of the global supply chain integration process, the core position of ports as logistics hubs has become increasingly prominent. According to the United Nations Conference on Trade and Development, approximately 80% of global trade volume is completed by sea transportation, and ports are the key nodes in this transportation chain [1]. As the world's largest trading nation, China completed 310 million standard container throughput at its national ports in 2023, with a year-on-year growth of 4.9% [2].

The continuous increase in container throughput has brought significant pressure to port operation scheduling: The traditional scheduling model relies on manual experience, which has problems such as information asymmetry, delayed decision-making, and unbalanced resource allocation, resulting in long vessel stay times, low equipment utilization, and high operating costs [3].

In this context, digitalization and intelligence have become the inevitable trend for port transformation. Digital twin technology, as a key technology to connect the physical world and the virtual space, can achieve real-time mapping of all elements of the operation scenario, multi-dimensional simulation and planning, and dynamic optimization decision-making, providing a new solution for port scheduling optimization [4]. This study focuses on the application mechanism of digital twin technology in port container operation scheduling, aiming to answer the core question of “How does digital twin technology drive the real-time scheduling optimization of port container operations and achieve the dual goals of efficiency improvement and cost control”? The research is mainly conducted from three aspects: technology empowerment, decision reconstruction, and performance optimization, and uses the grounded theory three-level coding method to systematically analyse multiple data sources such as participatory observation, system operation logs, and scheduling record reports. The study selects typical ports that have deployed digital twin systems as case objects, conducts field investigations, and forms triangulation verification through multiple data sources to ensure the credibility of the research conclusions

2. Literature Review and Theoretical Foundation

2.1 Digital Twin Technology and Its Application Research

The concept of digital twin was first proposed by Professor Grieves in 2003. Its core lies in achieving dynamic management of the system throughout its entire life cycle through the organic integration of physical entities, virtual models, and data links [3]. Tao et al. further proposed a five-dimensional structural model of digital twin, covering physical entities, virtual entities, connections, data, and services, providing a theoretical framework for the system construction of digital twin [5]. Tao Fei et al. conducted a systematic analysis and reflection on digital twin technology from a more macroscopic perspective [6]. In the port field, Neugebauer et al. conducted a systematic review of the application status, implementation paths, and decision-support value of digital twin in container terminal scheduling through global port case analysis [7], while Kastner et al. focused on the practical application and optimization path of digital twin technology in improving the operational efficiency of container terminals and carried out empirical research [8]. Zhang et al. explored the development direction of smart ports from the perspective of human-machine collaboration [9]. However, existing studies mostly focus on technical architecture design or optimization of a single link, lacking systematic examination of the entire scheduling process, and research on how digital twin is embedded in the scheduling decision-making process and its impact on the human-machine collaboration model is still relatively rare [8,9].

2.2 Research on Container Terminal Scheduling Optimization in Ports

Container terminal scheduling optimization has always been a research hotspot in the fields of logistics and operation management. Steenken et al. conducted a systematic review of the optimization problems of terminal operations, dividing the problems into berth allocation, crane scheduling, truck routing, and yard allocation, etc. [10]. Carlo et al. further analysed the coupling relationships among the sub-problems, indicating that integrated optimization is the key to improving overall efficiency [11]. In terms of scheduling methods, early studies mainly relied on mathematical programming and heuristic algorithms [12]. The relevant review studies have clearly demonstrated that the simulation optimization method has become the mainstream technical approach for the integrated planning of berth allocation and quay crane scheduling. [12]. In recent years, research has gradually evolved towards the integration of “technology + algorithm”, and relevant studies have pointed out that real-time data-driven and intelligent algorithm integration is the future development direction [9,13]. However, such research mostly remains at the level of theoretical modelling and lacks deep integration with actual operation scenarios, ignoring the collaborative relationship between human decision-making and the technical system.

2.3 Research on Human-Machine Collaborative Decision-Making

Human-machine collaborative decision-making is a core issue in the application of intelligent systems. Parasuraman et al. proposed a hierarchical model of human-machine function allocation, dividing human-machine responsibilities into four stages: information acquisition, analysis, decision-making, and action [13]. In the field of automated systems, studies have shown that as the level of automation increases, the role of humans shifts from operators to supervisors [9]. In the field of operation management, studies have shown that the key to human-machine collaborative decision-making lies in “mutual trust” and “explainability” - the system needs to be able to explain the logic of its recommended solutions, and operators need to understand the boundary of the system's capabilities [13].

2.4 Review of Research and Positioning of This Study

In summary, the existing research has the following gaps: First, the application research of digital twin in port scheduling mostly remains at the level of architecture design and simulation verification, lacking in-depth analysis of the interaction mechanism between technology, people, and performance in actual applications; second, there is relatively little research on how digital twin reshapes the scheduling decision-making process and affects the human-machine collaboration mode; third, there is a lack of field investigation and grounded theory analysis based on real scenarios, making it difficult to construct a process theory with explanatory power. This study intends to use a single-case exploratory design and the grounded theory analysis method to fill these research gaps.

3. Research Design

This case selects a core hub container port with a capacity of over 10 million TEUs in the domestic coastal area as the research object. This port implemented the world's first comprehensive digital twin technology platform for ports in 2022 and continuously iterated and upgraded the real-time scheduling system of digital twin container operations. The background of its construction, technical architecture, and application process all conform to the common characteristics of China's port digital transformation, and have typical research and reference value. The traditional scheduling mode of this port adopts an operation mode of “manual experience + local system assistance”, and has long faced common industry pain points such as information asymmetry, data silos, and delayed decision-making. There are problems such as mismatch of berth and crane resources, high empty driving rate of forklifts, and lack of forward-looking planning of the storage yard box positions. To solve these problems, the port built an operation real-time scheduling system based on digital twin technology, promoting the transformation of the scheduling mode from “manual experience-led” to “human-machine collaborative intelligent decision-making”.

3.1 Technology Empowerment

The digital twin container operation scheduling system of this port adopts a standardized hierarchical architecture of perception layer - transmission layer - modelling layer - application layer. In the perception layer, the system collects multi-source data such as vessel arrival dynamics, crane operation status, forklift positions, and storage yard box allocation through nodes such as Beidou positioning, equipment sensors, and video surveillance, achieving full-process real-time collection of data. This solves the problem of incomplete data collection and poor real-time performance in traditional scheduling. In the transmission layer and modelling layer, the system eliminates “data silos” through data standardization cleaning, heterogeneous data fusion, and real-time data graph construction. At the same time, it builds a high-precision virtual twin model that matches the physical port 1:1, achieving millisecond-level real-time mapping of physical operation scenarios and virtual models. In the application layer, the system develops five core functional modules: berth intelligent allocation, crane operation scheduling, forklift path optimization, storage yard box planning, and scenario simulation and planning, integrating the port's self-developed new generation container terminal control system, and possessing three core capabilities: full-process real-time mapping, multi-scenario simulation planning, and intelligent generation of scheduling plans.

3.2 Decision Reconfiguration

The implementation of the digital twin scheduling system has promoted the decision-making mode to transform from the traditional “single decision based on manual experience” to a hierarchical human-machine collaborative decision-making mode with different levels. In the early stage of system application, the decision-making mode was mainly “manual guidance + system information support”, and the dispatcher relied on their own experience to formulate plans. As the application deepened, the dispatcher's trust in the system gradually increased, forming a clear hierarchical operation logic: in routine operation scenarios, the system automatically generates a complete set of scheduling plans, and the dispatcher only needs to conduct compliance review and confirmation; in complex operation scenarios, the system simulates different plans through the virtual twin model, and the dispatcher combines experience for local optimization and adjustment, achieving “system scientific simulation + human experience judgment” for efficient collaboration; in emergency operation scenarios, the dispatcher initiates manual-led decision-making, and the system provides key data support in real time. After long-term integration, the dispatcher's role has transformed from “plan maker” to “plan reviewer and complex scenario optimizer”, and the system has taken on a large amount of standardized scheduling work, achieving complementary advantages between humans and machines.

3.3 Performance Optimization

The application of the digital twin scheduling system has promoted the efficiency improvement and cost control of container operations at the port. At the operation efficiency level, the system achieves a complete solution to the problem of information asymmetry through full-process real-time mapping and global data integration, effectively alleviating the phenomenon of resource mismatch. After the core intelligent terminal of the port implemented this system, the time for container pick-up and delivery in port has significantly decreased, the efficiency of ocean-going artery vessels when at berth has significantly improved, the operational continuity of core equipment such as cranes and forklifts has significantly enhanced, and the ineffective waiting time of equipment has been significantly reduced. In terms of operational costs, the ineffective operation rates of core equipment such as gantry cranes and container trucks have significantly decreased, effectively reducing the energy consumption per unit of operation for the equipment; the system has taken on a large amount of standardized scheduling tasks, significantly reducing the costs of manual communication and coordination in the scheduling process, and avoiding the problem of redundant manpower. From the perspective of the logic of collaborative optimization, the port has achieved global visualization and precise control of the entire operation process through digital twin technology, and optimized the operation process and reduced ineffective.

3.4 Data Analysis Process

This study follows the three-level coding path of grounded theory. In the open coding stage, the original data such as observation logs, system operation logs, and interview transcripts are analysed sentence by sentence to extract 57 initial concepts, which are then merged into 18 initial categories. In the main axis coding stage, the 18 initial categories are integrated into 5 core categories: technology empowerment, decision-making model, collaborative evolution, efficiency performance, and cost performance. In the selection coding stage, “digital twin-driven scheduling optimization” is selected as the core category, and the “process model of digital twin-driven scheduling optimization” is constructed. The analysis process is independently conducted by two members and compared regularly, and the saturation of the model is verified by the reserved original data.

4. Research Findings

The research found that the digital twin scheduling system adopts a hierarchical architecture, enabling real-time data collection and integration for the core operational processes. At the data collection level, the system achieves the fusion of multi-source heterogeneous data through standardized processing; at the real-time mapping level, the system builds a high-precision virtual port model to reflect the location, status, and operation progress of physical entities in real time; at the simulation and planning level, the system conducts scenario simulation and pre-planning based on real-time data. The integrity of data integration, the fidelity of

model mapping, and the timeliness of simulation and planning are the key elements determining the effectiveness of technology empowerment. After the system is deployed, a “system assistance - human-machine collaboration” decision-making model is formed: in routine scenarios, automated decision-making is carried out by the system, which automatically generates scheduling plans, and the dispatcher reviews and confirms; in complex scenarios, collaborative decision-making is conducted by the system, the simulation effect of the plan is considered by the dispatcher, and the decision is optimized based on experience; in emergency scenarios, manual-led decision-making is initiated by the dispatcher, and the system provides data support. As the system application deepens, the dispatcher's trust in the system gradually increases, the boundaries of human-machine responsibilities become increasingly clear, and a stable hierarchical decision-making model is formed. Based on the performance data analysis using the observation index system, the results show a significant collaborative optimization relationship between efficiency and cost. In terms of efficiency, the average operation time of ships is shortened, and the utilization rate of gantry cranes is improved; in terms of cost, the unit operation energy consumption of equipment decreases, and the efficiency of human resource allocation improves. The system optimizes the operation process, reduces equipment ineffective operation, improves efficiency while reducing energy consumption; through precise resource allocation, it shortens the operation response time and improves resource utilization, achieving a win-win situation of efficiency and cost.

5. Theoretical Model Construction

Based on the coding analysis of grounded theory, this study constructed the “process model of digital twin-driven port scheduling optimization”, revealing the complete transformation mechanism from technical input to performance output. Technological empowerment is the fundamental part of the model. The digital twin system, through real-time collection, integration of multi-source data and high-precision virtual simulation, provides the basic support for scheduling optimization. Data integration achieves transparency and visualization of the entire operation process, eliminating information asymmetry; simulation and deduction can predict the effects of different scheduling schemes in advance, reducing decision-making risks. Decision transformation is the core mechanism of the model. Technological empowerment realizes value transformation through decision model innovation, and the digital twin system changes the traditional single mode of manual decision-making, forming a human-machine collaborative decision-making mechanism that adapts to different scenarios: routine scenarios are automatically processed to improve efficiency, complex scenarios are jointly processed to balance scientific and flexibility, and emergency scenarios are led by humans to ensure safety. Performance output is the ultimate goal of the model. Human-machine collaborative decision-making ultimately promotes the coordinated optimization of efficiency and cost targets, by optimizing the operation process and rationally allocating resources, achieving the dual goals of efficiency improvement and cost reduction, forming a complete closed loop of “technology empowerment - decision optimization - performance improvement”.

6. Conclusion

The research results of this study are as follows: The digital twin technology realizes the real-time integration and virtual mapping of the entire process of port container operations through a hierarchical architecture, forming a hierarchical human-machine collaborative decision-making model, effectively promoting the collaborative optimization of operational efficiency improvement and cost reduction. Thus, the research conclusion drawn is that the complete transformation mechanism of digital twin-driven scheduling optimization can be summarized as a process model of “technology empowerment - decision transformation - performance output”, and the three elements form a closed-loop iterative optimization path. This study provides valuable reference significance for future research, mainly influencing the theoretical explanation of the application mechanism of digital twin in the port scheduling field, expanding the analysis framework of the human-machine collaborative decision-making model, and revealing the conditions for achieving dual-objective collaborative optimization of efficiency and cost. Future research should focus more on comparative studies of multiple cases to test the applicability and boundary conditions of the model, and conduct longitudinal tracking studies to observe the long-term evolution effect of technology application.

References

- [1] United Nations Conference on Trade and Development. (2023). *Review of maritime transport 2023: Towards a green and just transition*. United Nations. <https://unctad.org/publication/review-maritime-transport-2023>
- [2] Ministry of Transport of the People's Republic of China. (2024). Statistical bulletin on the development of the transportation industry in 2023. *China Water Transport*, (15), 31–35.
- [3] Grieves, M. (2014). *Digital twin: Manufacturing excellence through virtual factory replication* [White paper]. https://www.researchgate.net/publication/275211047_Digital_Twin_Manufacturing_Excellence_throug_h_Virtual_Factory_Replication (Original work published 2014)
- [4] Tao, F., Xiao, B., Qi, Q., Cheng, J., & Ji, P. (2022). Digital twin modeling. *Journal of Manufacturing Systems*, 64, 372–389.
- [5] Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., & Sui, F. (2018). Digital twin-driven product design, manufacturing and service with big data. *The International Journal of Advanced Manufacturing Technology*, 94(9), 3563–3576.
- [6] Tao, F., Zhang, H., Qi, Q., Zhang, M., Liu, W., Cheng, J., & Xue, R. (2020). Ten questions on digital twin: Analysis and reflection. *Computer Integrated Manufacturing Systems*, 26(1), 1–17.
- [7] Neugebauer, J., Heilig, L., & Voß, S. (2023). Digital twins in seaports: Current and future applications. In *International Conference on Computational Logistics* (pp. 202–218). Springer.
- [8] Kastner, M., Saporiti, N., Lange, A.-K., & Rossi, T. (2024). Insights into how to enhance container terminal operations with digital twins. *Computers*, 13(6), Article 138.
- [9] Zhang, R., Lv, J., Bao, J., & Zheng, Y. (2023). A digital twin-driven flexible scheduling method in a human–machine collaborative workshop based on hierarchical reinforcement learning. *Flexible Services and Manufacturing Journal*, 35(4), 1116–1138.
- [10] Steenken, D., Voß, S., & Stahlbock, R. (2004). Container terminal operation and operations research: A classification and literature review. *OR Spectrum*, 26(1), 3–49.
- [11] Makhado, N., Paepae, T., Sejeso, M., & Harley, C. (2025). Berth allocation and quay crane scheduling in port operations: A systematic review. *Journal of Marine Science and Engineering*, 13(7), Article 1339.
- [12] Li, B., Afkhami, P., Khayamim, R., Borowska-Stefańska, M., Wiśniewski, S., Fathollahi-Fard, A. M., Dulebenets, M. A., et al. (2025). An intelligent hyperheuristic algorithm for the berth allocation and scheduling problem at marine container terminals. *Transportation Research Part E: Logistics and Transportation Review*, 198, Article 104104.
- [13] Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 30(3), 286–297.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Research on the Brand Competitiveness Empowerment Path of Tencent's Silver-haired Service ESG Practices Driven by the Emotional Consumption of the Silver-haired Group in the Digital Era

Ke Mi*

School of Business, Xi'an International Studies University, Xi'an 710128, China

**Corresponding author: Ke Mi.*

Abstract

Against the dual background of a moderately aging society and digital consumption transformation, the consumption of the silver-haired group has shifted from function-oriented to emotion-oriented. Security anxiety, dignity pursuit, emotional loneliness, and communication barriers have become core emotional demands. Existing research insufficiently reveals the integrated mechanism among the silver-haired group's emotional consumption, ESG social dimension practices, and brand competitiveness, making it difficult to explain the value transformation logic of technology enterprises' silver-haired services. Taking Tencent's silver-haired service as a single case, this paper systematically analyzes the characteristics of the silver-haired group's emotional consumption, the ESG practice system, and the competitiveness empowerment path by adopting literature research, exploratory case study, and grounded theory methods, based on stakeholder theory, social exchange theory, and signaling theory. The findings show that: (1) The emotional consumption of the silver-haired group centers on security, respect, companionship, and smooth communication, characterized by emotion priority, low trial-and-error tolerance, high intergenerational transmission, and strong trust dependence; (2) Tencent has formed an ESG practice system responding to the emotional needs of the silver-haired group with the senior care exclusive hotline, AI invisible caregiver, silver-aged service base, and product-wide aging-friendly adaptation as the core; (3) Tencent's silver-haired service ESG practices transform into brand competitiveness through three paths: trust empowerment, word-of-mouth empowerment, and reputation empowerment; (4) Silver-haired ESG practices can achieve synergy between social responsibility and market benefits. This paper expands the theoretical boundary of emotional consumption, enriches the research on ESG value transformation mechanism, and provides a reference for technology enterprises to layout aging-friendly services and build brand competitiveness.

Keywords

silver-haired group, emotional consumption, ESG practices, silver-haired service, brand competitiveness, Tencent

1. Introduction

1.1 Research Background and Research Questions

China has entered a moderately aging society, with more than 310 million people aged 60 and above, and the number of elderly internet users reaching 161 million. The popularization of digital technology has highly onlineized the life, consumption, social interaction, and health management of the elderly. However, prominent problems such as operational obstacles, fraud risks, information discrimination, and lack of access to help have triggered strong emotional demands. The upgrading of consumption structure has driven the market from functional consumption to emotional consumption. The silver-haired group is willing to pay a premium for security, respect, companionship, and smooth communication, making emotional consumption a new growth point of the silver-haired economy.

Against this background, internet enterprises have accelerated the layout of aging-friendly transformation and silver-haired services, and incorporated them into ESG social dimension practices. As the earliest technology enterprise in China to layout silver-haired services, Tencent has formed a complete system covering online products, manual customer service, AI security, and offline bases, ranking among the top in the industry in terms of service volume and satisfaction. However, the theoretical circle has not yet answered: How does the silver-haired group's emotional consumption drive enterprises' ESG practices? Through what mechanism do silver-haired service-oriented ESG practices transform into brand competitiveness?

Existing research has three shortcomings: first, emotional consumption research is overly concentrated on young groups, with insufficient attention to silver-haired scenarios and rigid emotional demands; second, ESG research focuses on environmental and governance dimensions, with weak research on the value mechanism of social dimensions, especially silver-haired services; third, silver-haired service research mostly focuses on technical aging-friendliness, lacking integrated analysis with emotional demands, ESG strategies, and brand competitiveness. In view of this, this paper puts forward the following research questions: (1) What are the core demands and behavioral characteristics of the silver-haired group's emotional consumption in digital scenarios? (2) What is the system composition and demand response logic of Tencent's silver-haired service ESG practices? (3) What are the paths and mechanisms for Tencent's silver-haired service ESG practices to empower brand competitiveness? (4) What are the implications for technology enterprises to promote silver-haired ESG and brand upgrading?

1.2 Research Significance

Theoretical Significance: Firstly, expand the theory of emotional consumption to the silver-haired group, propose the concept of “rigid emotional demand of the silver-haired group”, and fill the gap in group coverage; secondly, construct an integrated framework of “emotional consumption—ESG response—brand competitiveness” to fill the theoretical context gap; thirdly, reveal the triple empowerment path of “trust—word-of-mouth—reputation” and open the black box of the mechanism for ESG social dimension to transform into brand value.

Practical Significance: Provide an operable path for technology enterprises to optimize silver-haired services, improve ESG efficiency, and build differentiated brand competitiveness; provide a “responsibility + profit” model for enterprises to enter the silver-haired economy; provide decision-making reference for building a digital society friendly to the elderly.

1.3 Research Content and Methods

Research Content: Define core concepts and theoretical basis; analyze the demands and characteristics of the silver-haired group's emotional consumption; sort out Tencent's silver-haired service ESG practice system; extract the triple empowerment path; form conclusions and implications.

Research Methods: (1) Literature Research Method: Systematically sort out literatures on emotional consumption, ESG, silver-haired economy, and brand competitiveness; (2) Exploratory Single Case Study: Follow the Eisenhardt and Yin paradigm, taking Tencent as a typical case; (3) Grounded Theory: Construct a theoretical model through open coding, axial coding, and selective coding; (4) Triangulation Verification: Cross-check with ESG reports, user comments, interviews, and media reports.

1.4 Innovations

(1) Innovation in Research Object: Incorporate the silver-haired group into the research framework of emotional consumption, making up for the deficiency of existing research biased towards young groups; (2) Innovation in Research Framework: Integrate emotional demands, ESG practices, and brand competitiveness, and establish a logical chain from micro-motivation to macro-strategy; (3) Innovation in Path Mechanism: Propose and verify the triple empowerment model of “trust—word-of-mouth—reputation”, and clarify the brand value transformation path of silver-haired ESG; (4) Innovation in Perspective: Start from rigid emotional demands, break through technology and function orientation, and fit the real decision-making logic of silver-haired consumption.

2. Literature Review and Theoretical Basis

2.1 Literature Review

2.1.1 Research on Emotional Consumption

Emotional consumption emphasizes that consumers conduct consumption behaviors to meet emotional needs. Holbrook & Hirschman proposed the paradigm of experiential consumption, and Wang Ning pointed out that digital consumption is turning to emotion orientation. Existing researches focus on Generation Z, self-pleasing consumption, live broadcast rewards, and fashion toy markets, lacking research on the emotional motivation, behavior, and decision-making of the silver-haired group.

2.1.2 Research on ESG and Enterprise Competitiveness

Most studies confirm that ESG is positively correlated with enterprise performance, but there are obvious biases: emphasis on environmental (E) and governance (G) dimensions, neglect of social (S) dimension; emphasis on overall correlation test, neglect of user psychology and emotional intermediary mechanism; emphasis on financial performance, neglect of the transformation of intangible assets such as brand trust, word-of-mouth, and reputation.

2.1.3 Research on Digital Consumption and Aging-friendly Services for the Silver-haired Group

Existing researches focus on digital divide, technology acceptance, anti-fraud security, and aging-friendly design. Studies by Shen Qi et al. pointed out that the elderly have strong emotional and companionship needs for smart products, but relevant researches have not yet risen to the theory of emotional consumption, nor combined with ESG and brand competitiveness.

2.1.4 Research on Tencent's Silver-haired Service and ESG

Existing achievements are mostly practical introductions and news reviews, lacking academic and mechanistic analysis of practices such as the senior care exclusive hotline, invisible caregiver, and silver-aged service base, especially failing to explain the transformation path of silver-haired services from social responsibility investment to brand competitiveness.

2.1.5 Literature Review

Existing researches are obviously fragmented: emotional consumption is separated from the silver-haired group; ESG practices are separated from emotional mechanisms; silver-haired services are separated from brand competitiveness. This paper aims to connect the logic of the three to form an integrated analysis framework.

2.2 Theoretical Basis

Stakeholder Theory: Enterprises can build long-term competitive advantages by meeting the demands of elderly users, families, society, and the government. **Social Exchange Theory:** Enterprises provide emotional value (security, respect, companionship), and users reward with trust, loyalty, and recommendation, forming a stable exchange and transforming into brand assets. **Signaling Theory:**

Enterprises send signals of responsibility, trustworthiness, and warmth to the market through silver-haired ESG practices, reducing information asymmetry and improving brand trust and reputation.

2.3 Definition of Core Concepts

Silver-haired Group: Users of digital products and services aged 60 and above. **Emotional Consumption of the Silver-haired Group:** Digital consumption choices and behaviors generated to meet emotional demands such as security, respect, companionship, and smooth communication. **ESG Practices for Silver-haired Service:** A series of social responsibility practices carried out by enterprises under the ESG social dimension, including aging-friendly transformation, protection of elderly rights and interests, digital inclusion, security protection, and emotional companionship. **Brand Competitiveness:** The sum of brand trust, word-of-mouth communication, social reputation, user stickiness, market expansion capability, and differentiated competitive advantages brought by silver-haired ESG practices.

3. Research Design

3.1 Case Selection

This paper selects Tencent's silver-haired service as the case based on: (1) Typicality: the earliest layout, the most complete system, the largest investment, and the widest coverage in China; (2) Enlightenment: serving more than 32 million person-times annually with 99% satisfaction, strong demonstration effect; (3) Data Availability: ESG reports, SSV reports, media reports, and user data are fully public.

3.2 Data Sources

Multi-source data triangulation is adopted to improve reliability and validity. **Secondary Data:** Tencent ESG reports (2020—2024), sustainable social value reports, official announcements, authoritative media reports, and industry data. **Primary Data:** Semi-structured interviews (elderly users, adult children users, customer service personnel, community workers); more than 300 online user comments; product experience and scene observation records.

3.3 Data Analysis Process

Follow the three-level coding process of grounded theory: (1) Open Coding: Label and conceptualize texts to form initial concepts; (2) Axial Coding: Summarize categories and logical relationships to form main categories; (3) Selective Coding: Refine core categories and construct a model of “emotional consumption → ESG response → competitiveness empowerment”; (4) Theoretical Saturation Test: Stop coding when no new categories and relationships appear.

4. Emotional Consumption of the Silver-haired Group: Demands and Characteristics

4.1 Core Emotional Demands of the Silver-haired Group

4.1.1 Security Anxiety

The elderly are highly sensitive to property risks, operational errors, and information leakage, fearing being deceived, misled, or deducted, and pursuing a traceable, confirmable, and accessible security experience.

4.1.2 Dignity Pursuit

The elderly are prone to frustration in complex operations, eager for patient, equal, clear, and non-discriminatory services, and refuse to be disliked, despised, or “labeled”.

4.1.3 Emotional Companionship

With the rising proportion of empty-nesters and solo-living elderly, the elderly seek social connection, emotional comfort, and value recognition through digital products, and companionship directly affects their willingness to use.

4.1.4 Smooth Communication

Dialect, hearing, and expression ability limitations make it difficult for the elderly to ask for help, requiring human services that can understand, answer patiently, assist, and are warm.

Table 1: Core Emotional Demands and Performance of the Silver-haired Group

Core Emotional Demands	Typical Performance	Impact on Consumption Decision
Security Anxiety	Fear of fraud, misoperation, information leakage, property loss	Prefer reliable, traceable platforms with customer service
Dignity Pursuit	Fear of discrimination, neglect, complex operations, misunderstanding	Value patient service, simple interface, equal communication
Emotional Companionship	Strong loneliness, narrow social circle, need for comfort	Prefer products with companionship and humanistic care
Smooth Communication	Dialect barriers, hearing decline, difficulty in expression	Rely on manual service, dialect support, one-to-one assistance

4.2 Behavioral Characteristics of Emotional Consumption of the Silver-haired Group

(1) Emotion is prior to function: security, convenience, and warmth are more important than technical parameters; (2) Low trial-and-error tolerance: a negative experience easily leads to long-term loss, and the cost of trust reconstruction is extremely high; (3) High intergenerational transmission: the evaluation of the elderly affects the choice of adult children, and the recommendation of adult children determines the acceptance of the elderly; (4) Strong trust and loyalty: high retention, repurchase, and recommendation after trust is established.

5. ESG Practice System and Demand Response Logic of Tencent's Silver-haired Service

5.1 Core ESG Practices of Tencent's Silver-haired Service

5.1.1 Senior Care Exclusive Hotline

Provide exclusive manual customer service, dialect support, patient response, and full-process assistance to solve difficulties in seeking help and communication barriers.

5.1.2 AI Invisible Caregiver

Realize fall detection, abnormal call for help, and emergency linkage based on vision and voice recognition, alleviating home security anxiety.

5.1.3 Silver-aged Service Base

Provide offline community teaching, one-on-one accompanying guidance, and social activities to improve dignity and integration.

5.1.4 Product-wide Aging-friendly Transformation

Launch senior mode on WeChat, QQ, payment, map, etc., with large fonts, simplified interfaces, voice guidance, and anti-misoperation.

5.1.5 Anti-fraud and Digital Inclusion for the Silver-haired Group

Strengthen security and fairness through anti-fraud science popularization, rural elderly assistance, public welfare training, and aging-friendly standard construction.

Table 2: Correspondence Between Tencent's ESG Practices and Emotional Demands

ESG Practice	Corresponding Emotional Demand	Value Orientation

Senior Care Exclusive Hotline	Communication barriers, dignity pursuit	Patient service, equal assistance
AI Invisible Caregiver	Security anxiety, health concerns	Safety protection, risk early warning
Silver-aged Service Base	Emotional companionship, digital isolation	Accompaniment, social connection, dignity
Product-wide Aging-friendly Adaptation	Dignity pursuit, operation anxiety	Simplified, low-threshold, user-friendly
Anti-fraud & Digital Inclusion	Security anxiety, information disadvantage	Risk prevention, equity, rights protection

5.2 Core Characteristics of Tencent's Silver-haired ESG Practices

(1) Centering on emotional demands, shifting from functional aging-friendliness to emotional aging-friendliness; (2) Full-scenario coverage of online products + offline services + AI security + manual companionship; (3) Perceptible, experienceable, and traceable practices that directly respond to emotional pain points; (4) Incorporated into the company's ESG and sustainable social value strategy with long-term investment.

5.3 Demand Response Logic

Security anxiety ← Invisible caregiver, anti-fraud mechanism, security risk control
Dignity pursuit ← Senior care exclusive hotline, patient service, aging-friendly simplification
Emotional companionship ← Silver-aged service base, social connection, offline activities
Communication barriers ← Dialect customer service, manual assistance, full-process operation on behalf of users

6. Triple Paths for Tencent's Silver-haired Service ESG Practices to Empower Brand Competitiveness

6.1 Trust Empowerment: Basic Level

Mechanism: ESG practices accurately respond to emotional pain points → meet emotional demands → user satisfaction → brand trust → user loyalty and repurchase. **Logic:** Security guarantee → peace of mind; patient service → warmth; easy to use → convenience; reliable solution → trust. **Results:** Improved satisfaction, increased retention, enhanced active choice, and reduced risk perception.

6.2 Word-of-mouth Empowerment: Extension Level

Mechanism: Positive experience of elderly users → intergenerational transmission in families → recognition and recommendation of adult children → user scale expansion and scenario penetration. **Logic:** Silver-haired experience → family word-of-mouth → social diffusion → new users and package transformation. **Results:** Significant intergenerational recommendation, growth of family users, and breakthrough of brand word-of-mouth.

6.3 Reputation Empowerment: Sublimation Level

Mechanism: Sustainable and systematic silver-haired ESG practices → social recognition and positive communication → improved brand reputation → policy support, industry demonstration, and capital favor. **Logic:** Responsibility investment → social identity → reputation capital → long-term brand barrier. **Results:** Improved ESG rating, selected into sustainability indices, increased policy cooperation, and solidified brand image of goodness.

6.4 Integrated Model of Triple Paths

Trust is the foundation, locking core silver-haired users; word-of-mouth is the diffusion, connecting family and intergenerational markets; reputation is the barrier, forming sustainable brand advantages. The three reinforce and synergize each other, jointly completing the transformation from social value to brand competitiveness.

Table 3: Triple Paths of ESG Practices Empowering Brand Competitiveness

Empowerment Path	Level	Core Transmission Mechanism	Final Effect
Trust Empowerment	Basic Level	Emotional satisfaction → Satisfaction → Trust → Loyalty	High retention, low churn, strong dependence
Word-of-mouth Empowerment	Extension Level	Senior experience → Intergenerational transmission → Family recognition	Customer expansion, brand penetration
Reputation Empowerment	Sublimation Level	ESG practice → Social recognition → Brand reputation	Long-term brand barrier, policy support

7. Research Conclusions and Management Implications

7.1 Research Conclusions

(1) Digital consumption of the silver-haired group has entered a stage driven by rigid emotional demands, with security, respect, companionship, and smooth communication as core demands; (2) Tencent's silver-haired service ESG practices are a systematic responsibility supply oriented to emotional demands, realizing the upgrade from functional aging-friendliness to emotional aging-friendliness; (3) Silver-haired service ESG practices transform into brand competitiveness through the triple path of “trust—word-of-mouth—reputation”, forming a replicable mechanism; (4) Silver-haired ESG practices are not cost investment, but an important strategic direction for technology enterprises to build differentiated brand assets.

7.2 Management Implications

(1) Anchor rigid emotional demands: take security, respect, companionship, and good communication as the core design principles of silver-haired products; (2) Promote the scenario-based implementation of ESG: transform responsibility into perceptible, experienceable, and measurable services; (3) Build a value closed loop: emotional satisfaction → trust → word-of-mouth → reputation → sustainable growth; (4) Adhere to long-termism: incorporate silver-haired services into strategies and ESG systems for continuous investment and iteration.

7.3 Research Limitations and Prospects

Limitations: Single case study, the universality of conclusions needs to be tested by multiple cases; quantitative data and large-sample empirical research can be further strengthened. **Prospects:** Carry out multi-case comparison; conduct questionnaire empirical testing; study the application of AI and large models in silver-haired emotional services; promote cross-cultural and cross-enterprise comparative research.

8. Conclusion

Against the backdrop of a moderately aging society and digital consumption upgrading, the emotional consumption of the silver-haired group has become a critical driver for enterprises to carry out ESG practices in elderly-oriented services. Taking Tencent's silver-haired service as a typical case, this study explores the formation mechanism of emotional consumption demands, the construction of ESG practice systems, and the triple empowerment paths of brand competitiveness through grounded theory and case analysis. The findings confirm that responding to the rigid emotional needs of the elderly—security, respect, companionship, and smooth communication—can effectively transform social responsibility investment into sustainable brand competitiveness through trust, word-of-mouth, and reputation empowerment. Such a model not only helps technology enterprises achieve the integration of social value and economic benefits but also provides a replicable framework for digital inclusion and age-friendly construction.

In the future, with the accelerated integration of AI, big models, and elderly-oriented services, the emotional interaction between enterprises and the silver-haired group will become more intelligent and precise. Subsequent research can expand to cross-enterprise comparisons, large-sample empirical tests, and cross-cultural contexts, so as to further improve the theoretical system of emotional consumption, ESG, and brand competitiveness in the silver-haired economy, and provide more systematic support for digital enterprises to fulfill social responsibilities and build long-term competitiveness.

References

- [1] Holbrook, M. B., & Hirschman, E. C. (1982). The experiential aspects of consumption: Consumer fantasies, feelings, and fun. *Journal of Consumer Research*, 9(2), 132–140.
- [2] Wang, N. (2021). Emotional turn in digital consumption and its sociological implications. *Sociological Studies*, 36(3), 21–38. (in Chinese)
- [3] Singh, S., & Verma, R. (2025). Building brand reputation and fostering customer loyalty through ESG practices: A strategic imperative for competitive advantage. In *ESG Frameworks for Sustainable Business Practices*. IGI Global.
- [4] Lin, J. (2024). ESG performance, corporate reputation, and brand valuation: Evidence from Chinese listed firms. *Journal of Business Ethics*, 187(3), 789–806.
- [5] Zou, Y. (2024). How ESG practices enhance brand equity: The mediating role of consumer trust. *Sustainability*, 16(12), 5123.
- [6] Kim, J., Kim, M., Im, S., & Choi, D. (2021). Competitiveness of e-commerce firms through ESG logistics. *Sustainability*, 13(20), 11548.
- [7] An, S., & Li, M. (2024). An SOR theory perspective on consumption alienation of silver consumers in emotional marketing within e-commerce live streaming. *ACM International Conference Proceedings Series*.
- [8] Jiang, Y., & Pan, C. (2025). Age-inclusive digital responsibility: A new form of responsibility in the resonance of aging and digitalization. *Business Ethics Quarterly*, 35(2), 1–26.
- [9] Ng, C. J., Teo, C. H., & Wattanapisit, A. (2026). Digital health for the ageing population. In *Handbook of Public Health in the Asia-Pacific*. Springer.
- [10] Wang, Z. Y. (2025). Research on the influence of emotional support on the purchase intention of the elderly in live e-commerce. *E-Commerce Letters*, 14(4), 1843–1849.
- [11] Zhang, Y., & Liu, H. (2025). Understanding the purchase decisions of silver consumers in short-form video platforms from the perspective of ERG needs. *Journal of Medical Internet Research*, 27(6), e10741116.
- [12] Tencent Holdings Ltd. (2024). *Tencent ESG Report 2023–2024*. Shenzhen: Tencent Sustainability Office.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

The Impact of Retirement on Urban Residents' Participation in Commercial Pension Insurance: An Empirical Analysis Based on Fuzzy Regression Discontinuity Design

Zilong Zhang*

School of Finance and Trade, Liaoning University, Shenyang 110000, China

**Corresponding author: Zilong Zhang.*

Abstract

As China's population aging accelerates, developing the third pillar of commercial pension insurance has become an important policy issue. However, the participation rate among urban residents remains persistently low. Existing studies mostly focus on static factors and overlook the dynamic impact of retirement—a critical life event. Based on data from the 2019 China Household Finance Survey (CHFS), this paper employs a fuzzy regression discontinuity design (FRDD), using the statutory retirement age as the policy cutoff, to precisely identify the causal effect of retirement on urban male residents aged 50–70 and their participation in commercial pension insurance. The study further examines the underlying mechanisms from three dimensions: income level, risk preference, and financial literacy. The results show that retirement significantly reduces the probability of urban residents participating in commercial pension insurance, decreasing the participation rate by 12.1 percentage points under the optimal bandwidth. Mechanism analysis reveals that retirement exerts a dual inhibitory effect by lowering residents' income levels and risk preferences. Although increased leisure time after retirement improves financial literacy and generates a certain positive effect, this positive influence is offset by the negative effects of income constraints and increased risk aversion. This paper provides rigorous empirical evidence on the causal relationship between retirement and commercial pension insurance participation, offering important policy implications for improving the third pillar of the pension system.

Keywords

retirement, commercial pension insurance, fuzzy regression discontinuity design, income level, risk preference, financial literacy

1. Introduction

With China entering a stage of moderate population aging, the degree of aging continues to deepen, and the societal pension burden is increasing day by day. How to provide for the elderly has become a major challenge facing Chinese society. The Fourth Plenary Session of the 20th Central Committee of the Communist Party of China, in conjunction with the 15th Five-Year Plan, proposed the integrated and coordinated development of the pension industry and the implementation of proactive national strategies to

address population aging. At present, among the “three pillars” of China’s pension system, basic social pension insurance covers the vast majority of retired individuals and holds an absolute advantage in the insurance market. However, in recent years, the overall fiscal expenditure on social security has been enormous and the burden has intensified, leading to severe challenges to the adequacy and sustainability of social pension insurance development. As the second pillar, enterprise annuities also face the problem of insufficient coverage, and the “five social insurances and two funds” represent a heavy contribution burden for enterprises that are currently undergoing tax and fee reductions. In this context, the third pillar—personal pensions represented by commercial pension insurance—has become the key to solving China’s current pension problems. The 2020 Central Economic Work Conference proposed to “regulate and develop the third pillar of pension insurance.” Commercial pension insurance has become an important channel for residents to open up new pension models and meet diversified pension needs.

Although policies have been actively promoted at the governmental level, it must be acknowledged that the participation rate of urban residents in commercial pension insurance in China remains at a relatively low level. The development speed of the pension finance market has not yet fully matched the severe reality of China’s population aging. Consequently, the academic community has conducted solid and multi-layered discussions on this issue. Most scholars have systematically and rigorously analyzed, from the perspectives of institutional design, market supply, and household characteristics, how factors such as severe homogenization of pension products, insufficient policy incentives [1], low resident income levels, and conservative risk preferences [2] constrain urban residents’ enthusiasm for participating in commercial pension insurance. More notably, some scholars have examined the actual impact of variables such as age and education level [3] and family demographic structure [4] on commercial insurance purchasing behavior. However, upon closer examination of the existing literature, it is not difficult to identify a significant and long-overlooked perspective: retirement, as a major life milestone, exerts a direct and fundamental exogenous shock on residents’ pension-related decisions. Retirement inevitably brings about fundamental changes in residents’ income structure and consumption patterns, and also alters their perception of their own health, thereby naturally influencing the intensity of their demand for pension security and their risk perception.

Statutory retirement can be regarded as an exogenous shock that has complex effects on commercial pension insurance participation. On the one hand, after retirement, residents’ income sources tend to become singular; the sharp decline in income level and the significant reduction in risk preference create a dual negative impact on commercial pension insurance participation. At the same time, retired residents’ income shifts to fixed sources such as pensions, and the decline in disposable income directly affects their economic ability to participate in insurance. Retirement also makes residents’ risk attitudes more conservative, significantly reducing their acceptance of long-term contribution financial products such as commercial pension insurance. On the other hand, the increase in leisure time among retired groups leads to greater interest in financial knowledge and related information, enhancing their understanding of commercial pension insurance products and potentially promoting participation. In addition, the change in social roles brought about by retirement may also influence residents’ information acquisition channels through social networks, indirectly affecting commercial pension insurance purchasing decisions.

Existing literature has not sufficiently discussed the actual effect of retirement on urban residents’ participation in commercial pension insurance. More importantly, it lacks rigorous identification of the causal relationship between the two. Traditional regression analysis methods struggle to adequately address sample selection bias and endogeneity issues, making estimation results susceptible to bias. Therefore, this paper makes full use of data from the China Household Finance Survey (CHFS) and adopts a fuzzy regression discontinuity design (FRDD) to conduct a solid and clear empirical analysis of the impact of retirement on urban residents’ participation in commercial pension insurance and its underlying mechanisms. Specifically, the fuzzy regression discontinuity design is a typical quasi-natural experiment method that uses retirement age as the policy cutoff. It can naturally and appropriately isolate the causal effect of retirement, making it a powerful econometric tool for identifying the true relationship between retirement and commercial pension insurance participation.

The marginal contributions of this paper are mainly reflected in the following three aspects: First, from the research perspective, this paper focuses on retirement as a key life event, breaking through the static analysis of individual characteristics in existing studies and enriching the research dimensions of factors

influencing commercial pension insurance participation. Second, in terms of research methodology, the use of fuzzy regression discontinuity design effectively addresses endogeneity issues and provides more rigorous empirical evidence for the causal relationship between retirement and commercial pension insurance participation. Third, at the policy level, the conclusions of this study can provide references for formulating differentiated strategies to promote commercial pension insurance and for improving the third pillar of the pension system.

The remainder of this paper is structured as follows: The second section reviews existing literature and proposes the theoretical foundation and research hypotheses. The third section introduces the data sources, research sample, variable measurement, and statistical model. The fourth section presents descriptive statistics, regression analysis, mediation effect tests, endogeneity treatment, robustness checks, and heterogeneity analysis. The final section offers conclusions and discussions.

2. Literature Review and Research Hypotheses

2.1 Literature Review

Against the backdrop of population aging, the construction of China's multi-pillar pension system has entered a critical stage. As the core of the third pillar, commercial pension insurance serves as an important support for alleviating the pressure on basic social pension insurance and meeting residents' diversified pension needs. However, according to data from the 2019 China Household Finance Survey (CHFS), the participation rate in commercial pension insurance among urban male residents aged 50–70 is only 7.5%, which severely constrains the functioning of the third pillar. Existing studies have mostly focused on static factors such as income, financial literacy, and traditional conceptions, while neglecting the dynamic impact of retirement—a key life event. Retirement not only signifies withdrawal from the labor market but also directly influences pension-related decision-making by reshaping income structure, risk preferences, time allocation, and information channels.

International research on the impact of retirement on individual financial decision-making started earlier and has become increasingly sophisticated in recent years in terms of research methods and mechanism identification. Early studies were largely based on the life-cycle theory, suggesting that increased income stability but limited growth potential after retirement would prompt individuals to reduce the proportion of risky assets in their portfolios. However, with the widespread application of quasi-natural experimental methods, scholars have achieved more rigorous identification of the causal effects of retirement. Shah and Flaherty [5] pointed out in their study that retirement, as an exogenous institutional cutoff, provides an ideal “quasi-experimental” setting for identifying behavioral differences in individuals' pension plan choices. Their research framework is consistent with the fuzzy regression discontinuity design adopted in this paper.

Regarding the dynamic relationship between retirement and risk preferences, scholars have found that retirement status significantly alters individuals' risk perception and asset allocation decisions. Lin et al. [6] constructed an optimal consumption and investment portfolio model for retirees under inflation and pension default risk, finding that retirees' degree of risk aversion is the core determinant of their asset allocation decisions. Particularly when facing longevity risk and market uncertainty, individuals exhibit a stronger conservative tendency. Wang and Jia [7] further deepened this line of research by analyzing equilibrium investment strategies in defined contribution pension plans under the assumption of stochastic risk aversion. They explicitly stated that ignoring the stochastic changes in risk preferences before and after retirement would lead to significant welfare losses. This finding provides solid theoretical support for understanding the decline in individuals' risk preferences after retirement and its inhibitory effect on pension participation.

In addition, existing studies have also deeply explored individuals' participation decisions in pension financial products from the perspective of risk-return trade-offs after retirement. Li and Wei [8] examined the optimal management strategies for defined contribution pension plans under the dual constraints of financial risk and mortality risk, emphasizing the trade-off process between fund security and long-term returns that individuals face after retirement. These studies indicate that retirement is not merely the termination of labor supply but a complex process that reshapes individuals' financial goals, risk tolerance, and decision-making capacity. Its ultimate impact on the demand for long-term financial products such as

commercial pension insurance depends on the combined effect of multiple factors, including income shocks, changes in risk preferences, and financial knowledge reserves.

Domestic research on the impact of retirement on financial behavior has mainly concentrated on two areas: asset allocation and consumption structure. Direct studies linking retirement with commercial pension insurance participation remain relatively scarce.

In terms of retirement and financial asset allocation, Jia Nan [9], using CHFS 2013 data and the FRDD method, found that male retirees do not change the allocation ratio between financial and non-financial assets but adjust the internal structure of financial assets—reducing the proportion of risky assets such as cash and stocks while increasing the proportion of low-risk assets such as bank deposits and wealth management products. The core mechanisms are the shift toward risk-averse attitudes, income shocks, and changes in financial knowledge caused by retirement. Feng Jin et al. [10] further revealed gender heterogeneity: after retirement, women, due to increased family caregiving responsibilities, show a stronger preference for liquid and low-risk financial assets.

In terms of retirement, consumption, and security needs, Zhao Xindong and Wang Hao [11], based on CHARLS data and using FRDD, found that China does not exhibit a universal “retirement consumption puzzle.” However, the proportion of enjoyment-oriented consumption such as entertainment and travel increases significantly, reflecting heightened attention to quality of life and risk protection after retirement. Wang Zhen et al. [12] discovered gender differences in post-retirement lifestyles: women utilize medical services to a greater extent and exhibit stronger health risk perception, providing evidence for gender differentiation in commercial pension insurance participation.

Regarding the influencing factors of commercial pension insurance participation, existing studies have mostly focused on static variables. Zheng Lu et al. [13] confirmed that traditional family concepts such as “raising children for old-age support” indirectly reduce urban residents’ participation rate by weakening financial trust and inhibiting financial literacy. Zhang Qiang et al. [14] found that income, education, and household wealth have significant positive effects on participation behavior. Chen Qifang [15] pointed out that rural residents’ awareness of commercial pension insurance and their attitude toward “raising children for old-age support” are key subjective factors.

Synthesizing the existing research findings, the academic community has conducted certain explorations into the influencing factors of commercial pension insurance participation and the economic and financial effects of retirement. However, targeted studies combining the two remain clearly insufficient, mainly manifested in three aspects: First, in terms of research perspective, most studies focus on the impact of static characteristic factors on commercial pension insurance participation, neglecting the shock effect of retirement as a dynamic life event, making it difficult to reveal the intrinsic relationship between retirement and commercial pension insurance participation. Second, in terms of mechanism analysis, the transmission pathways through which retirement affects commercial pension insurance participation have not been systematically and deeply explored. Existing studies sporadically mention the roles of income and risk preference but have not incorporated income level, risk preference, and financial literacy into a unified analytical framework, nor have they systematically tested the mediating effects of these three factors or identified their relative strengths using empirical data. Third, in terms of research methods and samples, there is a lack of precise causal identification and targeted sample selection. Most existing studies employ ordinary regression methods without adequately addressing endogeneity issues, and some fail to focus on groups near the retirement age, resulting in limited relevance and reliability of conclusions and making it difficult to accurately reflect the true impact of retirement on urban residents’ participation in commercial pension insurance.

2.2 Research Hypotheses

Retirement influences urban residents’ participation in commercial pension insurance through changes in core factors such as income status, risk perception, and financial knowledge reserves, operating at both direct and indirect levels. Drawing on existing theoretical and empirical logic, this paper first proposes the core hypothesis regarding the direct impact of retirement on urban residents’ participation in commercial pension insurance, and then puts forward corresponding mediating effect hypotheses centered on the three mechanisms of income level, risk preference, and financial literacy.

2.2.1 Core Hypothesis: The Direct Impact of Retirement on Urban Residents' Participation in Commercial Pension Insurance

Retirement marks the end of most residents' professional careers and directly alters their economic status, pension expectations, and financial decision-making objectives. In terms of participation ability, post-retirement wage income declines significantly. Even with pension income, the overall reduction in disposable income lowers residents' ability to pay for commercial pension insurance. In terms of participation willingness, after retirement, residents' life focus shifts toward the actual realization of pension security, moving from the "accumulation period" to the "consumption period" in pension planning, which reduces their willingness to participate in commercial pension insurance. In terms of institutional substitution, after retirement, residents begin to receive benefits from basic social pension insurance, and some treat it as the primary source of pension security, creating a substitution effect on commercial pension insurance.

Considering the actual situation of urban residents in China, post-retirement commercial pension insurance participation is subject to multiple negative influences from the above factors. Based on this, the paper proposes the core hypothesis:

H1: Retirement significantly reduces the probability of urban residents participating in commercial pension insurance.

2.2.2 Mediating Effect Hypothesis of Income Level

Income level is a core economic factor affecting residents' participation in commercial pension insurance and serves as an important transmission channel for the impact of retirement. On the one hand, there is a direct causal relationship between retirement and residents' income level. Retirement means the termination of wage income, while pension income is typically lower than pre-retirement wage income, leading to a significant decline in overall income level—this is the most direct economic shock of retirement. On the other hand, income level directly determines residents' ability to purchase pension financial products. As a non-essential financial consumption, participation in commercial pension insurance highly depends on disposable income. A decline in income level directly reduces residents' payment capacity, thereby inhibiting participation.

Thus, income level may play a mediating role between retirement and urban residents' participation in commercial pension insurance. That is, retirement indirectly suppresses participation by reducing residents' income level. Based on this, the paper proposes the following hypotheses:

H2.a: Retirement significantly reduces urban residents' income level.

H2.b: Income level plays a mediating role in the impact of retirement on urban residents' participation in commercial pension insurance; that is, retirement inhibits participation by lowering residents' income level.

2.2.3 Mediating Effect Hypothesis of Risk Preference

Risk preference is an important subjective characteristic in residents' financial decision-making and changes significantly across different life stages. It is also a key transmission mechanism through which retirement affects commercial pension insurance participation. On the one hand, retirement significantly alters residents' risk preference. During employment, residents are in the wealth accumulation stage with stable wage income support, resulting in stronger risk-bearing capacity and relatively higher risk preference. After retirement, income sources become more stable but decline overall, wealth accumulation enters a stagnation phase, and the rigid demand for pension security leads residents to focus more on wealth preservation than appreciation. Consequently, risk aversion increases and risk preference decreases significantly. On the other hand, risk preference is positively correlated with commercial pension insurance participation. As a financial product combining security and investment attributes, participation in commercial pension insurance requires a certain level of risk acceptance. Individuals with higher risk preferences are more willing to engage with it, while a decline in risk preference reduces participation willingness.

Therefore, risk preference exerts a mediating effect between retirement and urban residents' participation in commercial pension insurance. That is, retirement indirectly inhibits participation by lowering residents' risk preference. Based on this, the paper proposes the following hypotheses:

H3.a: Retirement significantly reduces urban residents' risk preference.

H3.b: Risk preference plays a mediating role in the impact of retirement on urban residents' participation in commercial pension insurance; that is, retirement inhibits participation by lowering residents' risk preference.

2.2.4 Mediating Effect Hypothesis of Financial Literacy

Financial literacy represents important human capital for residents to understand and participate in the pension finance market, and its mechanism between retirement and commercial pension insurance participation is distinctive. On the one hand, after retirement, residents experience a substantial increase in leisure time, allowing them more energy to focus on and learn financial knowledge. At the same time, to better preserve and increase pension wealth, residents actively enhance their financial knowledge reserves. Therefore, retirement may significantly improve urban residents' financial literacy. On the other hand, the level of financial literacy is significantly positively correlated with commercial pension insurance participation. Higher financial literacy helps residents better understand product features, protection value, and operational mechanisms of commercial pension insurance, reducing decision-making costs caused by information asymmetry and thereby increasing participation probability.

Thus, financial literacy plays a special mediating role between retirement and urban residents' participation in commercial pension insurance. That is, retirement indirectly promotes participation by enhancing residents' financial literacy. This mediating effect partially offsets the negative impact of retirement on commercial pension insurance participation, forming a suppression (masking) effect. Based on this, the paper proposes the following hypotheses:

H4.a: Retirement significantly improves urban residents' financial literacy.

H4.b: Financial literacy plays a mediating role in the impact of retirement on urban residents' participation in commercial pension insurance; that is, retirement promotes participation by enhancing residents' financial literacy, thereby partially offsetting the negative impact of retirement.

3. Model Specification

Commercial pension insurance is an important tool for improving the quality of retirement and managing risks. Although commercial pension insurance serves as a significant supplement to China's pension security system, the participation rate in commercial pension insurance in China has remained at a relatively low level for a long time. Jia Nan (2022) pointed out that retirement, as a major exogenous shock, causes significant changes in the composition of household risky assets, with households showing a stronger preference for investing in low-risk financial products. As an important component of low-risk financial products, the impact of retirement on participation in commercial pension insurance can be divided into two aspects: On the one hand, the economic shock caused by retirement may lead to a substantial decline in disposable income, thereby reducing households' participation rate in commercial pension insurance. On the other hand, retirement increases residents' leisure time, which improves their ability to acquire financial knowledge and may stimulate their participation in commercial pension insurance. Therefore, the actual net effect of retirement on the participation rate in commercial pension insurance remains to be empirically tested.

The household decision to participate in commercial pension insurance can be characterized using a discrete choice model. Let $busi\_insure$ denote household participation in pension insurance and $Retire$ indicate whether the household head is retired. The impact of retirement on whether a household holds commercial pension insurance can be estimated using Equation (1):

$$busi\_insure_i = \alpha + \beta \cdot retire_i + \gamma \cdot X_i + \varepsilon_i \quad (1)$$

where X_i represents control variables, including household characteristics, individual characteristics, and regional characteristics; ε_i is the error term; β captures the effect of retirement on household participation in commercial pension insurance; and γ denotes the coefficients of the control variables.

In the above model specification, the parameters may suffer from endogeneity bias. Certain unobservable omitted variables may simultaneously affect both retirement behavior and commercial pension insurance participation, such as health status and job characteristics. Traditional OLS regression has difficulty isolating the interference of these factors, making the estimation results prone to bias. At the same time, age is highly correlated with retirement, so direct regression cannot effectively distinguish the age effect from the retirement effect on commercial pension insurance participation.

Drawing on the approach of Jia Nan (2020) and other studies examining the economic effects of retirement, this paper adopts the fuzzy regression discontinuity design (FRDD) as the core research method. Its core logic lies in the fact that China's statutory retirement age policy creates an exogenous policy cutoff, which allows the exogenous shock of retirement behavior to be separated from the continuous effect of age, thereby achieving precise identification of the causal effect of retirement on commercial pension insurance participation. China implements a statutory retirement age system. Relevant policies stipulate that the normal retirement age for men working in government agencies, public institutions, state-owned enterprises, and collective enterprises is 60 years. The retirement ages for female workers and female cadres are 50 and 55 years, respectively. However, some workers and cadres may retire early or delay retirement due to special circumstances. Therefore, the decision of whether to retire is not determined solely by the running variable "age" but is also influenced by other unobservable factors, such as individual health status, position, and job nature. Overall, the probability of retirement for individuals above the statutory retirement age is substantially higher than for those below it. Hence, this constitutes a classic fuzzy regression discontinuity design.

In this study, the running variable is age (Age), the treatment variable is retirement status (Retire), and the outcome variable is participation in commercial pension insurance (Busi_insure). The constant c represents the cutoff value of age. $Retire_i = 1$ if $Age_i \geq c$, and $Retire_i = 0$ if $Age_i < c$. According to Lee and Lemieux (2010), the local average treatment effect (LATE) of the fuzzy regression discontinuity design can be expressed as Equation (2):

$$LATE \varPi E[Y_1 - Y_0 | age = c] = \frac{\lim_{\varepsilon \downarrow 0} E(busi\_insure | age = c + \varepsilon) - \lim_{\varepsilon \uparrow 0} E(busi\_insure | age = c - \varepsilon)}{\lim_{\varepsilon \downarrow 0} E(retire | age = c + \varepsilon) - \lim_{\varepsilon \uparrow 0} E(retire | age = c - \varepsilon)} \quad (2)$$

To estimate the above local average treatment effect, this paper employs local linear regression to fit the samples near the cutoff and constructs the following model:

$$busi\_insure_i = \alpha + \beta_1 age_i^* + \beta_2 D_i + \beta_3 D_i \times age_i^* + \mu_i \quad (3)$$

where $busi\_insure_i$ is a binary variable indicating whether individual i participates in commercial pension insurance (1 = participates, 0 = otherwise); age_i^* is the centered age variable; D_i is the age dummy variable, which equals 1 if $age_i \geq 60$ and 0 otherwise. Its coefficient β_2 reflects the average treatment effect of retirement on commercial pension insurance participation. The interaction term $D_i \times age_i^*$ controls for the difference in the marginal effects of age on commercial pension insurance participation on either side of the cutoff. μ_i is the random error term.

In actual estimation, weighted local linear regression is applied to the above model, with weights determined by the kernel function $K(\cdot)$. The core idea is to assign higher weights to samples near the cutoff and lower weights to samples farther away. The parameter estimates are obtained by minimizing the weighted sum of squared residuals. The objective function is:

$$\min_{\alpha, \beta_1, \beta_2, \beta_3} \sum_{i=1}^n K\left(\frac{age_i^*}{h}\right) [busi\_insure_i - \alpha - \beta_1 age_i^* - \beta_2 D_i - \beta_3 D_i \times age_i^*]^2 \quad (4)$$

where h is the bandwidth, reflecting the range of samples near the cutoff included in the estimation, and $K\left(\frac{age_i^*}{h}\right)$ is the kernel function used to calculate sample weights.

Following Imbens and Kalyanaraman (2012), this paper uses the IK algorithm to determine the optimal bandwidth. This algorithm achieves optimal bandwidth selection by minimizing the mean squared error

(MSE) of the regression function at the cutoff, effectively balancing estimation bias and variance. The baseline model adopts the triangular kernel function, which is the most widely used in regression discontinuity designs and provides a more reasonable weight distribution for samples near the cutoff. To test the robustness of the results, subsequent estimations will also employ quadratic and uniform kernel functions, and report results under 0.5 times and 2 times the optimal bandwidth.

4. Data, Variables, and Model Specification

4.1 Data Sources and Sample Construction

The core data used in this paper come from the 2019 China Household Finance Survey (CHFS). The survey employs a stratified, three-stage, probability proportional to size (PPS) sampling design, covering 29 provinces (autonomous regions and municipalities) across China. It provides a comprehensive, systematic, and high-quality reflection of Chinese households' assets, liabilities, income, consumption, and demographic characteristics, offering a solid micro-level data foundation for studying household financial behavior.

To precisely identify the causal effect of retirement as an exogenous shock on commercial pension insurance participation, this study applied strict screening and sample construction to the raw data. The specific steps are as follows: First, to align with the research theme and eliminate interference from urban-rural institutional differences, only samples with urban household registration (*hukou*) were retained. Second, to clearly capture the policy cutoff effect at the statutory retirement age of 60 and following common practice in relevant studies, the sample was restricted to male urban residents aged 50–70. This choice is based on two main considerations: (1) the statutory retirement age for urban male employees in China is primarily 60 years, providing a clear and single cutoff point, which avoids the identification complications caused by multiple retirement ages for women (50 for female workers and 55 for female cadres); (2) focusing on individuals near the retirement age allows for more effective control of pure life-cycle age effects, making the identified “jump” more likely to stem from the transition in retirement status rather than natural aging. Third, to eliminate the “age heaping” effect at the cutoff—where individuals exactly aged 60 may be in a transitional retirement period and their financial decisions mix pre- and post-retirement information—this paper excluded samples with an exact age of 60, a treatment consistent with mainstream literature. Fourth, to ensure estimation validity, observations with missing values in key variables were deleted. Finally, to avoid distortion of non-parametric estimation results by extreme values, all continuous variables were winsorized at the 1% and 99% percentiles. After the above rigorous data processing procedures, a balanced sample containing 6,214 valid observations was obtained, providing data support for subsequent fuzzy regression discontinuity analysis.

4.1.1 Variable Definitions and Descriptive Statistics

The variable system of this paper is constructed around the core dependent variable, key treatment variable, mechanism variables, and a series of control variables. Detailed definitions and measurement methods are presented in Table 4-1.

Commercial pension insurance participation (*busi_insure*): This is the core outcome variable of this study. It is defined based on whether the household owns any commercial pension insurance products according to the CHFS questionnaire. If the household holds any type of commercial pension insurance (including annuity insurance, endowment insurance, or other insurance products with pension savings functions), it is assigned a value of 1; otherwise, 0. This variable directly measures households' market-based pension reserve behavior beyond formal social pension insurance.

Retirement status (*retire*): This is the core treatment variable. It is defined as a complete transition in work status. Specifically, if the respondent is currently “not working” and the reason is “retired, early retired, or internally retired,” the individual is considered retired and assigned a value of 1. If the respondent is currently engaged in any income-generating work (including employed, self-employed, or farming), they are considered not retired and assigned a value of 0. This definition captures the essential feature of retirement as an interruption in labor market participation.

4.1.2 Mechanism Variables

To deeply explore the transmission channels through which retirement affects commercial pension insurance participation, this paper constructs three potential mediating variables:

Income level (wage): Measured by the natural logarithm of the individual's annual wage income. Retirement typically implies a shift from stable labor income to relatively fixed pension income, which is expected to generate a significant negative income shock.

Risk preference (risk): Constructed based on the CHFS survey item regarding investment risk attitudes. Individuals who choose "high risk, high return" or "slightly higher risk, slightly higher return" are defined as risk-preferring and assigned a value of 1. Those who choose "average risk, average return," "slightly lower risk, slightly lower return," or "unwilling to take any risk" are defined as risk-averse and assigned a value of 0. Theoretically, risk-taking willingness is expected to decline after retirement.

Financial literacy (finance): Following the construction method of Yin Zhichao et al. (2014), this variable is generated by summing the number of correct answers to three questions on interest rate calculation, inflation understanding, and investment risk perception. The value ranges from 0 to 3, with higher values indicating richer financial knowledge. Increased leisure time after retirement may promote the learning and accumulation of financial knowledge.

4.1.3 Control Variables

To control for other factors that may influence commercial pension insurance participation, this paper introduces two categories of control variables: individual characteristics and household characteristics. Individual characteristic variables include years of education (edu) (converted from highest educational attainment), marital status (marriage) (married = 1, unmarried = 0), and self-assessed health status (health) (1–5 scale, higher values indicate better health). Household characteristic variables include family size (familysize), annual household income (income) (natural logarithm), total household assets (asset) (natural logarithm), and homeownership (house) (owns = 1, otherwise = 0). These variables control for households' resource constraints, wealth levels, and basic security conditions.

Table 1: Variable Definitions

Variable Type	Variable Name	Symbol	Variable Definition
Core Variables	Commercial pension insurance participation	busi_insure	1 if participates in commercial pension insurance, 0 otherwise
	Retirement status	retire	1 if not working and retired/early retired, 0 otherwise
	Age	age	Continuous variable
Mechanism Variables	Income level	wage	Natural log of wage income
	Risk preference	risk	1 if prefers high or slightly higher risk, 0 otherwise
	Financial literacy	finance	1–5 scale, higher values indicate greater financial knowledge
Individual Controls	Years of education	edu	Converted from education level: 0 (no schooling), 6 (primary), 9 (junior high), 12 (senior high), 13 (technical secondary), 15 (junior college), 16 (bachelor), 19 (master), 22 (PhD)
	Marital status	marriage	1 if married, 0 if unmarried, divorced, or widowed
	Health status	health	1–5 scale, higher values indicate better health
Household Controls	Family size	familysize	Total number of family members
	Household income	income	Natural log of annual household income
	Household assets	asset	Natural log of total household assets
	Homeownership	house	1 if owns home, 0 otherwise

*Note: Full sample N=6,214; left-side sample: ages 50–59; right-side sample: ages 61–70. **

Table 2 presents the descriptive statistics, which preliminarily reveal the sample characteristics and differences on either side of the cutoff. First, the retirement rate exhibits the expected sharp jump at the cutoff, surging from 4.1% on the left side to 65.5% on the right side. This provides evidence of the strong enforcement of China's statutory retirement age policy and serves as a key prerequisite for the validity of the subsequent fuzzy regression discontinuity design. Second, the overall participation rate in commercial pension insurance is 7.5%, indicating that, as a market-based pension tool, its penetration among urban households still needs improvement. More importantly, the participation rate is 10.5% on the left side of the

age-60 cutoff (ages 50–59, predominantly working) and drops sharply to 3.0% on the right side (ages 61–70, predominantly retired), intuitively showing a significant downward trend. The mean values of other control variables are generally similar on both sides of the cutoff, preliminarily satisfying the “local randomization” assumption of regression discontinuity design. However, strict validation through subsequent continuity tests is still required.

Table 2: Descriptive Statistics

Variable Name	Symbol	Full Sample	Left Side (50–59)		Right Side (61–70)	
		Mean	Mean	SD	Mean	SD
Commercial pension insurance participation	busi_insure	0.075	0.105	0.307	0.03	0.17
Retirement status	retire	0.288	0.041	0.199	0.655	0.476
Age	age	58.715	54.356	2.781	65.205	2.817
Years of education	edu	10.028	10.084	3.23	9.944	3.58
Marital status	marriage	0.937	0.939	0.24	0.934	0.249
Health status	health	3.277	3.334	0.974	3.192	0.962
Family size	familysize	2.732	2.882	1.038	2.507	0.984
Household income	income	10.687	10.583	2.275	10.843	1.585
Household assets	asset	13.273	13.242	1.483	13.32	1.555
Homeownership	house	0.915	0.92	0.272	0.908	0.289

5. Empirical Results and Analysis

5.1 Baseline Regression Results: The Causal Effect of Retirement on Commercial Pension Insurance Participation

Table 3 reports the local average treatment effect (LATE) of retirement on commercial pension insurance participation estimated using the fuzzy regression discontinuity design. We successively present the estimation results under 0.5 times the optimal bandwidth, the optimal bandwidth (IK bandwidth), and 2 times the optimal bandwidth to examine the robustness of the conclusions across different bandwidth selections.

Table 3: The Impact of Retirement on Commercial Pension Insurance Participation: Baseline Regression Results

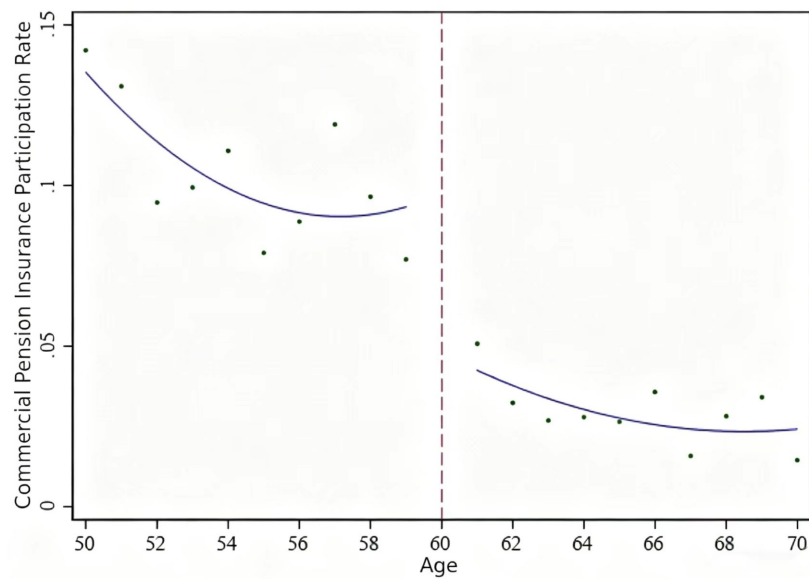
	(1)	(2)	(3)
	0.5× Bandwidth	Optimal Bandwidth	2× Bandwidth
Retirement Shock	-0.052	-0.121***	-0.099***
	(0.081)	(0.045)	(0.036)
Covariates	No	No	No

*Note: *, **, and *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Estimations use the triangular kernel function without additional covariates.*

Although the negative impact of retirement on the participation rate is not statistically significant under 0.5 times the optimal bandwidth, under both the optimal bandwidth and 2 times the optimal bandwidth, the estimated coefficients on the retirement variable are significantly negative at the 1% statistical level. Specifically, under the optimal bandwidth, retirement leads to an average significant decline of 12.1 percentage points in the probability of urban male residents participating in commercial pension insurance. Given that the overall participation rate in commercial pension insurance in the sample is only 7.5%, this magnitude of decline is economically substantial, implying that the retirement shock reduces the participation probability by more than 160% in relative terms. The insignificance of the coefficient under 0.5 times the optimal bandwidth is typically due to the sharp reduction in sample size caused by the excessively narrow bandwidth, which increases estimation variance—a common phenomenon in regression discontinuity analysis and does not undermine the core conclusions drawn from the optimal bandwidth.

Figure 1 provides a more intuitive illustration of this causal effect. The horizontal axis represents age centered at 60 years, and the vertical axis represents the local mean of the commercial pension insurance participation rate. It is clearly visible that on the right side of the cutoff (at 0), the fitted curve exhibits a distinct downward jump. This visual “discontinuity” precisely reflects the local average treatment effect induced by the exogenous shock of retirement, corroborating the numerical results in Table 3.

Figure 1: Relationship Between Age and Commercial Pension Insurance Participation

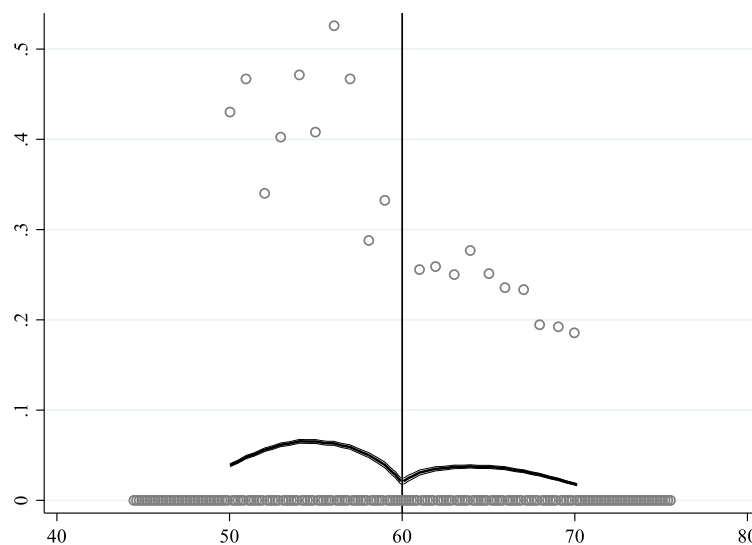


5.2 Robustness Tests

To ensure the reliability of the baseline regression results, we conduct a systematic and rigorous set of robustness tests.

The primary prerequisite for the validity of a regression discontinuity design is the absence of manipulation of the running variable (age) near the cutoff. If individuals could precisely choose to retire around age 60, it would undermine the “local randomness” near the cutoff. We plot the density distribution of sample age centered at 60 (Figure 2). The results show that the density function curve is smooth and continuous at the cutoff, with no obvious bunching or dips. Further, the McCrary test statistic is 0.044 with a standard error of 0.122, failing to reject the null hypothesis that the density function is continuous at the cutoff. This indicates that there is no systematic manipulation of birth timing or retirement timing by individuals in the sample to obtain retirement benefits. The distribution of the running variable is exogenous, satisfying the identification assumptions.

Figure 2: Changes in the Density Distribution of Sample Age Before and After the Cutoff



5.3 Sensitivity Analysis with Different Kernel Functions and Bandwidths

We test the robustness of the results by changing the kernel function specification in the estimation. As shown in Table 4, after replacing the baseline triangular kernel with quadratic and uniform kernels, the estimated coefficients on retirement remain significantly negative at the 1% level under both the optimal bandwidth and 2 times the optimal bandwidth, with magnitudes highly consistent with the baseline results. This indicates that our core finding—that retirement significantly reduces commercial pension insurance participation—does not depend on the choice of a specific kernel function, demonstrating high robustness of the conclusion.

Table 4: Robustness Tests with Alternative Kernel Functions

	Quadratic Kernel Function			Uniform Kernel Function		
	0.5× BW	Optimal BW	2× BW	0.5× BW	Optimal BW	2× BW
Retirement Shock	-0.064 (0.078)	-0.124*** (0.043)	-0.096*** (0.035)	-0.149** (0.067)	-0.093*** (0.034)	-0.093*** (0.034)
Covariates	No	No	No	No	No	No

Note: Robust standard errors are in parentheses. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

5.4 Continuity Tests of Covariates and Conditional Independence

This paper conducts a systematic and rigorous test of the smoothness of all selected control variables at the cutoff by using each covariate as the outcome variable in a regression discontinuity estimation and observing the coefficient at the cutoff. According to theoretical logic, if the treatment and control groups are comparable in other characteristics at the cutoff, all coefficients should be insignificant. As clearly shown in Table 5, the coefficients on marital status, health status, family size, household assets, and homeownership are all insignificant. In clear contrast, years of education (edu) and household total income (income) both exhibit highly significant jumps at the cutoff. A reasonable explanation naturally follows: there may exist some selection in the sample, such that individuals with higher education levels and higher household income are more likely to choose early or delayed retirement, leading to systematic differences in these two variables on either side of the policy cutoff.

Table 5: Continuity Tests of Covariates at the Cutoff

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	edu	marriage	health	familysize	income	asset	house
Retirement Shock	0.422** (0.203)	-0.013 (0.016)	0.032 (0.064)	-0.045 (0.065)	0.296** (0.117)	0.046 (0.096)	-0.019 (0.018)

Note: *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

To ensure that the estimation results are not contaminated by variables that violate the continuity assumption, in subsequent regressions that include covariates, we exclude 'edu' and 'income' and retain only the other smooth covariates.

5.5 Regression Results with Smooth Covariates Included

Since this paper first controls for various smooth covariates near the cutoff (marital status, health, family size, household assets, and homeownership), we re-estimate the model accordingly. The results are presented in Table 6. They are highly consistent with the baseline results: whether under the optimal bandwidth or 2 times the optimal bandwidth, the negative effect of retirement remains highly significant. Therefore, we can naturally and appropriately conclude that retirement has a clear inhibitory effect on commercial pension insurance participation.

Table 6: Regression Results with Covariates Included

	(1)	(2)	(3)
	0.5× BW	Optimal BW	2× BW
Retirement Shock	-0.045 (0.082)	-0.127*** (0.046)	-0.110*** (0.036)
Covariates	Yes	Yes	Yes

Note: Control variables include marital status, health, family size, household assets, and homeownership. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

5.6 Placebo Tests

To demonstrate that the significance of the baseline results is not due to chance or specific model specifications, we conduct placebo tests. Specifically, we construct two fictitious cutoffs where no retirement policy shock should exist—age 55 and age 65—and repeat the baseline model estimation at these hypothetical cutoffs. The results in Table 7 show that, whether using age 55 or 65 as the fictitious cutoff, the estimated coefficients on “retirement” become insignificant across all bandwidth settings, and their magnitudes change dramatically compared with the baseline results (even turning positive in some cases). This strongly confirms that the significant effect observed at the true age-60 cutoff is indeed caused by the statutory retirement policy shock at age 60, rather than being a spurious result driven by the data or model.

Table 7: Placebo Tests: Changing the Cutoff Position

	Age 55 as Cutoff			Age 65 as Cutoff		
	0.5× BW	Optimal BW	2× BW	0.5× BW	Optimal BW	2× BW
Retirement Shock	2.142	-0.129	-0.242	0.339	-0.272	0.161
	(3.657)	(0.307)	(0.519)	(0.910)	(0.183)	(0.225)
Covariates	No	No	No	No	No	No

Note: Robust standard errors are in parentheses. All estimated coefficients are insignificant.

5.7 Preliminary Discussion of Heterogeneity Analysis

Although the local average treatment effect identified by the fuzzy regression discontinuity design has excellent internal validity, a full and rigorous examination of the complexity of retirement’s impact naturally requires heterogeneity analysis of the results (note that strict subsample regression discontinuity designs impose high requirements on sample size; thus, this section serves as an extended discussion). From a theoretical perspective, there are clear and logically supported conjectures: the impact of the retirement shock is likely a function of heterogeneity in household financial literacy and initial wealth levels. Specifically, households with richer financial knowledge are more aware of the importance of long-term pension planning and are therefore less affected by short-term income shocks in their insurance decisions. Meanwhile, households with greater assets have stronger capacity to withstand the decline in retirement income, resulting in more stable participation rates in commercial pension insurance. Therefore, subsequent research should systematically examine these heterogeneity channels using grouped regressions or by introducing interaction terms.

6. Mechanism Tests

The baseline regression has confirmed the significant inhibitory effect of retirement on commercial pension insurance participation. The question then arises: through which specific channels does this effect operate? This section examines the mechanisms through three pathways: income shock, change in risk preference, and change in financial literacy. Following the fuzzy regression discontinuity framework, we estimate each of the three mechanism variables as the outcome variable.

Mechanism 1: Negative Income Shock from Retirement

The most direct and fundamental economic consequence of retirement is a sharp change in income sources and levels. Individuals shift from relying on labor wages to depending on pensions, usually accompanied by a sudden drop in disposable income. This negative income shock may suppress the demand for commercial pension insurance in two ways: first, the “budget constraint effect,” whereby the household’s cash flow available for insurance premiums decreases; second, the “priority reallocation effect,” whereby, after income declines, households prioritize limited resources for basic living consumption, thereby crowding out investment in long-term pension savings.

The results in Table 8 provide clear and strong evidence for the discussed mechanism. When the natural logarithm of individual wage income is used as the dependent variable, retirement yields statistically highly significant negative coefficients under 0.5 times, the optimal, and 2 times the optimal bandwidth. Specifically, under the optimal bandwidth, retirement reduces individual wage income by an average of 7.3 log units, meaning actual income declines by approximately 98%. Therefore, the post-retirement income

reality naturally becomes the fundamental constraint causing retired households to cut spending on non-essential, long-term financial products.

Table 8: Mechanism Test: The Impact of Retirement on Income Level

	(1)	(2)	(3)
	0.5× BW	Optimal BW	2× BW
Retirement Shock	-7.920***	-7.322***	-7.074***
	(1.476)	(0.789)	(0.621)
Covariates	No	No	No

Mechanism 2: Decline in Risk Preference Induced by Retirement

Life-cycle theory offers a clear and logically rigorous view: as individuals age and income sources become more fixed, risk aversion naturally increases. Retirement, as the endpoint of one's career, both fixes future income streams and substantially alters individuals' perceptions of future uncertainty. Consequently, financial decisions after retirement generally become more conservative. In line with this, although commercial pension insurance is relatively stable, it is undeniably still a financial contract with evident risks (such as insurer operational risk and inflation risk) and relatively low liquidity. Therefore, when risk preference declines, households' willingness to hold such assets correspondingly decreases.

The results in Table 9 rigorously confirm the existence of this mechanism: retirement has a statistically highly significant negative impact on the risk preference indicator. Under the optimal bandwidth, retirement reduces the probability of individuals exhibiting risk-preferring behavior by 9.0 percentage points. It can thus be naturally and appropriately concluded that individuals indeed become more risk-averse after retirement. This also naturally leads to changes in asset allocation: a greater willingness to hold risk-free or low-risk assets such as cash and bank deposits, while reducing allocation to higher-uncertainty financial assets such as commercial pension insurance.

Table 9: Mechanism Test: The Impact of Retirement on Risk Preference

	(1)	(2)	(3)
	0.5× BW	Optimal BW	2× BW
Retirement Shock	-0.136***	-0.090***	-0.075***
	(0.051)	(0.027)	(0.022)
Covariates	No	No	No

Mechanism 3: Changes in Financial Literacy Accompanying Retirement

From the current mechanism tests, an interesting yet somewhat contradictory phenomenon emerges quite naturally: Table 10 clearly shows that retirement has a distinct positive effect on financial literacy. Under the optimal bandwidth, retirement increases individuals' financial literacy score by an average of 0.86 points (out of a full score of 3). This paper offers the following explanation: the "leisure substitution effect." After retirement, individuals are freed from busy work and have more leisure time to pay attention to financial information and learn financial knowledge, which is also conducive to managing their savings and pensions.

However, the impact of this mechanism has two clearly logical sides. In theory, higher financial literacy helps individuals understand the importance of pension planning and the functions of insurance products, which should facilitate participation. Yet, when combined with the first two mechanisms, it can be naturally and appropriately observed that the "promoting effect" of increased financial literacy is actually completely overwhelmed by the more prominent and direct negative effects of income constraints and risk aversion. Therefore, for newly retired groups, the immediate financial realities and psychological adjustments remain the primary factors dominating their decisions. Even if they objectively possess more financial knowledge, it is extremely difficult for them to translate it into actual long-term pension investment actions. In other words, this reflects the typical "knowing is easy, doing is difficult" problem regarding financial literacy among China's retired population. It also suggests that existing commercial pension insurance products have not yet adequately met the specific expectations of post-retirement groups regarding liquidity, security, and returns.

Synthesizing the tests of the above three mechanisms, we can outline a clear pathway diagram of how retirement influences commercial pension insurance decisions. As an exogenous shock, retirement first directly generates a severe negative income shock, tightening household budget constraints. Second, it induces a systematic decline in risk preference, altering households' psychological thresholds for asset

allocation. These two factors form a powerful combined force that suppresses demand. Although retirement may incidentally produce a modest increase in financial literacy through increased leisure time, this positive effect is completely offset under the dominance of the dual negative effects of income and risk. Ultimately, the net effect manifests as a substantial decline in the probability of commercial pension insurance participation.

Table 10: Mechanism Test: The Impact of Retirement on Financial Literacy

	(1)	(2)	(3)
	0.5× BW	Optimal BW	2× BW
Retirement Shock	0.862***	0.859***	0.898***
	(0.311)	(0.170)	(0.136)
Covariates	No	No	No

7. Conclusion

Based on data from the 2019 China Household Finance Survey and employing a fuzzy regression discontinuity design, this paper precisely identifies the causal effect of retirement on commercial pension insurance participation among urban male residents aged 50–70 and examines the roles of three mediating mechanisms: income level, risk preference, and financial literacy. The core conclusion is that retirement exerts a significant negative impact on urban residents' participation in commercial pension insurance. Under the optimal bandwidth, retirement reduces the participation probability of urban male residents by a significant 12.1 percentage points. This result remains robust after various robustness checks.

From the mechanism perspective, the impact of retirement on commercial pension insurance participation can be analyzed clearly and hierarchically: retirement leads to a sharp decline in wage income and a marked reduction in risk preference levels. Thus, retirement itself constitutes the dual core driving force inhibiting commercial pension insurance participation. Although the increased leisure time brought by retirement raises residents' financial literacy levels and theoretically has a potential promoting effect on participation, this positive effect is completely offset by the dual negative effects of income constraints and risk aversion. The final outcome is still a decline in participation probability. Therefore, this paper breaks through the static analytical framework of influencing factors on commercial pension insurance participation in existing literature. By adopting a quasi-natural experimental method, it properly addresses endogeneity issues, clarifies the intrinsic logic between retirement and participation behavior, and provides highly valuable empirical insights for the construction of the third pillar of the pension system.

This study has certain limitations: since the sample is restricted to urban males aged 50–70, it does not include females, rural residents, or different occupational groups; thus, the generalizability of the conclusions is somewhat limited. Additionally, because cross-sectional data are used, the study can only estimate the short-term impact of retirement and cannot directly examine its dynamic effects. Moreover, the measurement of commercial pension insurance participation is limited to whether or not one participates and does not consider deeper participation indicators such as premium expenditure or product types. Future research should expand the sample scope, adopt panel data, and improve the measurement dimensions.

Based on the research conclusions, the following targeted policy recommendations are proposed to increase the participation rate of retired groups in commercial pension insurance and improve the multi-pillar pension security system: First, implement differentiated policy incentives for retired groups, optimize measures such as tax-deferred pension insurance and premium subsidies to alleviate their income constraints. Second, innovate commercial pension insurance products tailored to the low-risk preferences of retired groups, focusing on developing low-risk, stable-return, and liquidity-matched products to lower participation thresholds. Third, leverage the foundation of improved financial literacy among retired groups to carry out targeted pension financial education, address the “knowing is easy, doing is difficult” problem, and promote the transformation of financial knowledge into actual participation behavior. Fourth, provide early pension planning education for working groups to guide them to participate during the wealth accumulation stage and avoid the income and psychological shocks after retirement.

In summary, retirement is the most direct and significant exogenous shock affecting urban residents' pension-related financial decisions and can also be regarded as the core factor constraining the development of the third pillar of pension insurance. Therefore, in the context of deepening population aging, to activate

the commercial pension insurance market, it is essential first to clarify the financial characteristics, psychological preferences, and pension needs of retired groups. Only by addressing the practical constraints on retired groups' participation through policy incentives, product innovation, and financial education can commercial pension insurance form genuine synergies with basic social pension insurance and enterprise annuities, thereby systematically and solidly building China's multi-pillar pension security system.

References

- [1] Dong, K. Y., & Zhang, D. (2017). China's pension finance: Practical dilemmas, international experiences, and response strategies. *Administrative Management Reform*, (08), 16-21. <https://doi.org/10.14150/j.cnki.1674-7453.2017.08.004>.
- [2] Ba, S. S., & Huang, K. H. (2025). Evaluation of Chinese residents' retirement preparation and the impact of delayed retirement. *Population and Economics*, (01), 1-19.
- [3] Li, D., Ding, J. S., & Ma, S. (2019). The impact of social interaction on household commercial insurance participation: An empirical analysis based on data from the China Household Finance Survey (CHFS). *Journal of Financial Research*, (07), 96-114.
- [4] Fan, G. Z., & Wang, H. Y. (2015). Family demographic structure and household demand for commercial life insurance: An empirical study based on data from the China Household Finance Survey (CHFS). *Journal of Financial Research*, (07), 170-189.
- [5] Shah G. G. ,Flaherty C. M. (2013). Incorporating employee heterogeneity into default rules for retirement plan selection. *The Journal of Human Resources*, 48(1), 198-235.
- [6] Lin, Z., Lai, C., & Li, R. (2025). Optimal consumption and portfolio selection for retirees under inflation and pension default risk. *North American Journal of Economics and Finance*, 79102446-102446. DOI:10.1016/J.NAJEF.2025.102446.
- [7] Wang, L., & Jia, B. (2025). Equilibrium investment strategies for a defined contribution pension plan with random risk aversion. *Insurance: Mathematics and Economics*, 125103140-103140. DOI:10.1016/J.INSMATHECO.2025.103140.
- [8] Li, W., & Wei, P. (2024). Optimal defined-contribution pension management with financial and mortality risks. *ASTIN Bulletin: The Journal of the IAA*, 54(3), 621-658.
- [9] Jia, N. (2020). The impact of retirement on urban household financial asset allocation in the context of aging: Evidence from a fuzzy regression discontinuity design. *Statistical Research*, 37(04), 46-58. <https://doi.org/10.19343/j.cnki.11-1302/c.2020.04.004>.
- [10] Feng, J., & Han, X. (2017). The impact of retirement age policy on family caregiving and labor force participation. *The World Economy*, 40(06), 145-166. <https://doi.org/10.19985/j.cnki.cassjwe.2017.06.008>.
- [11] Zhao, X. D., & Wang, H. (2018). The impact of retirement on household consumption: Evidence from a fuzzy regression discontinuity design. *Wuhan University Journal (Philosophy & Social Sciences)*, 71 (01), 167-174. <https://doi.org/10.14086/j.cnki.wujss.2018.01.015>.
- [12] Wang, Z., & Zeng, Z. (2022). The impact of retirement on the utilization of medical and health services: An empirical study based on regression discontinuity. *Modern Preventive Medicine*, 49(16), 2992-2997. <https://doi.org/10.20043/j.cnki.MPM.202202325>.
- [13] Zheng, L., & Xu, M. X. (2021). Do traditional family values inhibit urban residents' participation in commercial pension insurance? An empirical analysis from the perspectives of financial trust and financial literacy. *Journal of Financial Research*, (06), 133-151.
- [14] Zhang, Q., & Yang, Y. (2017). Analysis of influencing factors on participation in commercial pension insurance. *Journal of Huazhong Agricultural University (Social Sciences Edition)*, (05), 138-143+150. <https://doi.org/10.13300/j.cnki.hnwkxb.2017.05.018>.

- [15] Chen, Q. F. (2016). Analysis of influencing factors on rural residents' willingness to purchase commercial pension insurance. *Theory and Practice of Finance and Economics*, 37(01), 59-63+109. <https://doi.org/10.16339/j.cnki.hdxbcjb.2016.01.009>.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

How Does Green Innovation Affect Corporate Profitability? An Empirical Study Based on the Mediating Effect of Green Patents of Agricultural Listed Companies

Ziqi Zhang^{†,*} and Shuangxiao Li[†]

School of Economics and Management, Shandong Agricultural University, Tai'an 271000, China

[†]*These authors contributed equally to this work and share first authorship*

^{*}*Corresponding author: Ziqi Zhang.*

Abstract

In the context of global sustainable development strategies and China's "dual carbon" goals, green innovation has become a critical driving force for the transformation and development of agricultural enterprises. This study takes Chinese A-share agricultural listed companies from 2016 to 2024 as the research sample and systematically examines the impact mechanism of green innovation on corporate profitability. It particularly introduces green patents as a mediating variable to reveal the transmission path between green innovation and profitability. The study employs the entropy weight method to construct a comprehensive corporate profitability score. Green innovation is measured from two dimensions: R&D intensity and financing constraints. Empirical tests are conducted using multiple regression models. The results show that R&D intensity has a significant negative impact on corporate profitability, while financing constraints exert a significant positive impact. Green patents play a positive mediating role between R&D intensity and profitability, but a negative mediating role between financing constraints and profitability. Heterogeneity analysis indicates that the above relationships are more pronounced in the northern regions and in the post-epidemic period. The findings provide empirical support and policy implications for agricultural listed companies to optimize green innovation resource allocation and enhance profitability, and offer experience and path references from the Chinese capital market for global agricultural green transformation and sustainable development.

Keywords

green innovation, corporate profitability, green patents, financing constraints, agricultural enterprises

1. Introduction

The United Nations 2030 Agenda for Sustainable Development lists green development as a core global issue. The World Economic Forum continues to promote corporate climate action, and the International Organization for Standardization (ISO) also guides corporate transformation through environmental management standards. Under this international consensus, China has proposed the "dual carbon" goals,

incorporating carbon peaking and carbon neutrality into the overall layout of ecological civilization construction. Facing increasingly stringent environmental regulations and rising social awareness of environmental protection, green innovation has become an important strategic means for enterprises to achieve sustainable development. Green innovation refers to the introduction of new practices in product design, production processes, or management methods that reduce negative environmental impacts and improve resource efficiency [1]. Enterprises are using this as a core strategy to build sustainable competitive advantages [2].

For agricultural listed companies, green innovation not only involves technological innovations in production such as water-saving irrigation, precision fertilization, and green prevention and control of pests and diseases, but also encompasses the transformation of business models such as circular agriculture, ecological planting, and low-carbon breeding. These innovations hold special significance for the control of agricultural non-point source pollution and the intensive utilization of resources. The above characteristics of agricultural green innovation make the relationship between environmental benefits and financial performance particularly worthy of attention. On the one hand, based on the natural resource-based view and the Porter Hypothesis, green innovation can ultimately be transformed into financial gains by reducing energy consumption and compliance costs and by opening up green product markets [3]. Recent large-sample studies have also confirmed a significant and robust positive relationship between green innovation and corporate profitability [4]. On the other hand, some studies point out that green innovation often requires substantial upfront investment and has a long payback period, and its effects are highly dependent on internal and external contextual factors of the enterprise [5]. Therefore, an in-depth exploration of the relationship between green innovation and corporate profitability holds important theoretical value and practical significance for agricultural listed companies in optimizing green strategies and achieving a win-win outcome of economic and environmental benefits.

Green patents, as the most direct intellectual property output of corporate green innovation activities, play a key mediating role between green innovation and corporate profitability [7,8]. On the one hand, obtaining green patents requires enterprises to invest substantial R&D resources, and the application and authorization process itself reflects the substantive efforts of corporate green innovation [7]. On the other hand, the technological value embedded in green patents can be transformed into corporate financial gains through channels such as efficiency improvement and cost reduction [6,9]. However, existing studies still lack systematic examination of the mediating transmission mechanism of green patents. How factors such as the pollution intensity of the industry in which the enterprise operates, its own financial flexibility, and its sustainable development performance affect the mediating effect of green patents remains to be further revealed [8,9]. In addition, the impact of green patents on corporate profitability may exhibit non-linear characteristics; that is, R&D investment in the short term may squeeze corporate profits, while long-term accumulation is expected to achieve positive returns [8,10]. Therefore, this study takes green patents as a mediating variable and systematically examines its transmission mechanism in the process of green innovation affecting corporate profitability, in order to provide a more detailed theoretical explanation and empirical evidence for understanding the complex relationship between the two.

In summary, this paper takes Chinese A-share agricultural listed companies as the research object. Based on panel data from 2016 to 2024, it constructs multiple regression models to systematically examine the impact of green innovation on corporate profitability and introduces green patents as a mediating variable to reveal its internal transmission mechanism. In terms of measuring green innovation, this paper analyzes it from two dimensions: R&D intensity and financing constraints. The former reflects the level of corporate resource investment in green technologies, while the latter characterizes the degree of funding constraints faced by enterprises in carrying out green innovation. Furthermore, from the perspectives of regional heterogeneity and temporal heterogeneity, this paper examines the impact of institutional differences between northern and southern regions and the shock of the COVID-19 pandemic on the main effects, with a view to revealing the differentiated performance of corporate green innovation under changes in the external environment. It aims to provide theoretical basis and practical reference for enterprises to optimize resource allocation and achieve sustainable development in the process of green transformation. It also holds important practical significance for promoting the control of agricultural non-point source pollution, improving the level of intensive resource utilization, and achieving the long-term strategic goal of high-quality agricultural development.

2. Theoretical Basis and Hypotheses

2.1 Green Innovation and Corporate Profitability

R&D intensity, as a core proxy variable for corporate innovation input, is widely used to characterize the level of resource investment in enterprises' green innovation [14]. This indicator is typically measured as the ratio of R&D expenditure to sales revenue and can effectively reflect the degree of emphasis enterprises place on green technology R&D and their resource allocation orientation [15]. Green R&D activities inherently feature large-scale investment, long payback periods, and high technological uncertainty. In the short term, high R&D expenditures may crowd out funds for daily operations. In addition, the time lag between technological breakthroughs and commercial application means that R&D intensity is difficult to translate into profitability contributions in the current period and may even exert a negative squeezing effect on profitability [5,10]. For agricultural listed companies, agricultural R&D activities are highly dependent on uncontrollable factors such as natural conditions and biological cycles. The process from laboratory to field application, and from experimental demonstration to large-scale adoption, often requires crossing longer validation cycles and higher adoption thresholds [44]. This implies that R&D investment in agricultural enterprises is even more difficult to achieve results conversion in the short term. Therefore, this paper chooses to measure corporate green innovation from the dimension of R&D intensity, so as to depict the level of resource investment and innovation efforts of agricultural listed companies in the green transformation process.

Accordingly, the following hypothesis is proposed:

H1a: R&D intensity has a negative impact on corporate profitability.

Financing constraints, as an important dimension for measuring corporate green innovation, have received extensive attention in the existing literature. Hottenrott and Peters argue that financing constraints reflect an enterprise's ability to raise external funds for innovation activities and constitute a key dimension for measuring the efficiency of innovation resource acquisition and allocation [19]. Some studies suggest that moderate financing constraints may force enterprises to optimize resource allocation and improve the efficiency of R&D investment [45]. Specifically, financing constraints can impose financial discipline on corporate management, suppress excessive investment and agency problems, and encourage enterprises to allocate limited capital more prudently, thereby enhancing capital utilization efficiency [45]. For agricultural listed companies, agricultural enterprises generally face structural financing difficulties such as a high proportion of fixed assets, unstable collateral values, and mismatch between production cycles and capital recovery cycles, resulting in relatively limited external financing channels [44]. In this context, financing constraints are not only a restriction on external fund acquisition for agricultural enterprises but also constitute an endogenous resource allocation constraint mechanism. Moderate financial pressure may compel agricultural enterprises to evaluate green investments more cautiously, prioritize projects with shorter payback periods and relatively mature technologies, and thereby improve the allocation efficiency of limited funds. Therefore, this paper chooses to measure corporate green innovation from the dimension of financing constraints, so as to depict the funding constraints faced by agricultural listed companies in the green transformation process and empirically test its impact on profitability.

Accordingly, the following hypothesis is proposed:

H1b: Financing constraints have a positive impact on corporate profitability.

2.2 The Mediating Role of Green Patents

Green patents, as the core intellectual property output of corporate green innovation activities, play an important transmitting role between green innovation and corporate profitability [21,22]. For agricultural listed companies, green patents hold particularly special significance. Agricultural technologies are characterized by highly context-specific applications, long commercialization cycles, and deep integration with natural resource endowments. These features make the acquisition and utilization of green patents a critical link connecting R&D investment and corporate profitability. Existing studies still lack systematic examination of the mediating transmission mechanism of green patents, and their specific pathways and boundary conditions remain to be further revealed [23,24]. Based on this, the present study takes green

patents as a mediating variable and systematically examines its mediating effect in the processes through which R&D intensity and financing constraints respectively influence corporate profitability.

R&D intensity reflects the level of resource investment by enterprises in green technological innovation and serves as the fundamental guarantee for carrying out green innovation activities. Enterprises with higher R&D intensity are more likely to produce high-quality green technological achievements, which need to be protected through patent applications and then commercialized [21]. R&D investment has a significant positive impact on green patent output, while the accumulation of green patents can enhance enterprises' market competitiveness and profitability through channels such as technology licensing and product differentiation [22]. From the perspective of the underlying mechanism, R&D intensity first promotes substantive breakthroughs in green technology through capital investment and talent allocation; the key output of this process is green patents. As the legal embodiment of corporate technological capabilities, green patents not only prevent knowledge spillovers from being freely exploited by competitors, but also signal the enterprise's genuine capabilities in green technology to external investors, thereby alleviating information asymmetry [23]. On this basis, enterprises can obtain technology revenues through patent licensing or enhance product pricing power through differentiation advantages formed by patent protection, ultimately converting innovation inputs into economic returns [46]. This series of linkages reveals that the impact of R&D intensity on profitability is not entirely direct, but is partially transmitted through the intermediate link of green patents [47].

Accordingly, the following hypothesis is proposed:

H2a: Green patents play a mediating role in the relationship between R&D intensity and corporate profitability.

Financing constraints reflect the degree of funding restrictions faced by enterprises in carrying out green innovation activities. When financing is constrained, enterprises' ability to engage in green innovation is restricted, and the output of green patents may also decrease accordingly. However, green patents, as signals of corporate innovation capability, can convey competitive advantages in green technology to the market, helping to alleviate information asymmetry and attract external financing [22]. The acquisition of green patents can, to a certain extent, break through the constraints imposed by financing restrictions on corporate innovation activities: on the one hand, enterprises possessing green patents are more likely to gain recognition from external investors and broaden their financing channels; on the other hand, improved financing conditions in turn provide further financial support for enterprises to carry out green innovation, forming a virtuous cycle. On this basis, the accumulation of green patents can improve enterprises' financing conditions, reduce financing costs, and thereby create greater profit margins for enterprises [23].

Accordingly, the following hypothesis is proposed:

H2b: Green patents play a mediating role in the relationship between financing constraints and corporate profitability.

3. Research Design

3.1 Variable Measurement

Table 1 reports the definitions and measurement methods of the main variables used in this study. The dependent variable is the corporate comprehensive profitability score, the independent variables are R&D intensity and financing constraints, the mediating variable is green patents, and company size, listing age, and the proportion of independent directors are selected as control variables.

Table 1: Variable Measurement

Variable Type	Variable Name	Symbol	Description and Calculation
Dependent Variable	Corporate Comprehensive Profitability Score	Score1	Comprehensive score calculated by objectively weighting seven financial indicators—current ratio, quick ratio, debt-to-asset ratio, accounts receivable turnover, total asset turnover, return on equity (ROE), and gross profit margin—using the entropy weight method.
Independent Variables	R&D Intensity	Research	R&D expenses / Total expenses
	Financing	WW	Financing constraints measured by the WW index constructed by

Variable Type	Variable Name	Symbol	Description and Calculation
	Constraints		Whited and Wu. It is a linearly weighted index based on six indicators: corporate cash flow, dividend payout, leverage level, firm size, industry sales growth rate, and the firm's own sales growth rate.
Mediating Variable	Green Patents	EP	Natural logarithm of (1 + total number of invention patents, utility model patents, and design patents) divided by $\ln(1 + \text{R\&D expenditure})$
Control Variables	Company Size	Size	Natural logarithm of total assets at the end of the year
	Listing Age	Age	Natural logarithm of (listing years + 1)
	Proportion of Independent Directors	Indep	Number of independent directors / Total number of board members

(1) Dependent Variable

The dependent variable in this study is the corporate comprehensive profitability score. Existing studies have mostly measured corporate profitability from a single dimension, such as solvency, operational capability, or profitability. Compared with a single financial indicator, the comprehensive profitability score can more fully reflect the overall financial condition and operating results of an enterprise [25]. As a multidimensional comprehensive evaluation indicator, the comprehensive score contains richer financial information and can effectively measure the overall financial performance of the enterprise. Referring to the methods of existing studies [25], this paper selects seven financial indicators—current ratio, quick ratio, debt-to-asset ratio, accounts receivable turnover, total asset turnover, return on equity (ROE), and gross profit margin—and applies the entropy weight method for objective weighting [26] to obtain the corporate comprehensive profitability score.

(2) Independent Variables

One of the independent variables in this study is R&D intensity. R&D intensity is an important way for enterprises to acquire core competitiveness and innovation advantages. Especially for knowledge-intensive enterprises, R&D activities constitute the primary means of technological innovation [27]. Existing studies have mostly measured R&D intensity as the ratio of R&D expenditure to operating revenue [28-30], while some scholars directly use the amount of R&D investment [31]. For agricultural enterprises, R&D activities are characterized by long cycles, high risks, and strong influence from natural conditions; thus, the measurement of R&D investment needs to take into account the overall operating scale of the enterprise. Following the approach of existing studies [27], this paper adopts the ratio of R&D expenses to total expenses as the measure of R&D intensity for agricultural listed companies. A higher ratio indicates a greater relative intensity of R&D investment, which can better reflect the degree of resource allocation toward technological innovation by agricultural enterprises.

The other independent variable in this study is financing constraints. Financing constraints refer to the significant difference between the costs of internal and external financing, which makes it difficult for enterprises to raise funds to support their investment activities [32]. Existing studies mainly adopt multi-indicator composite index methods to measure financing constraints, among which the KZ index [33] and the WW index [18] are two widely used indicators in academia [34,35]. Compared with the KZ index, the WW index is constructed based on a structural investment model and includes multiple dimensions such as corporate cash flow, dividend payout, long-term debt, firm size, industry sales growth, and firm sales growth. It can more comprehensively reflect the degree of external financing constraints faced by enterprises [18]. Therefore, this paper uses the WW index as the measure of financing constraints for agricultural listed companies. A higher WW index indicates a higher degree of financing constraints faced by the enterprise.

(3) Mediating Variable

The mediating variable in this study is green patents. Green patents are an important output of corporate green innovation activities and reflect an enterprise's technological innovation capability in the fields of environmental protection and sustainable development [22,36]. For agricultural listed companies, the measurement of green patents is particularly critical because the application of agricultural technologies has strong regional dependence and ecological sensitivity. Green patents not only embody technological

breakthroughs but also reflect the enterprise's adaptability to regional resource endowments and environmental protection requirements. Existing studies mainly use the number of green patent grants or applications as proxy variables [22,36]. Considering the differences in firm size and innovation investment levels, simply using the number of green patents cannot accurately measure green innovation efficiency. Referring to the approach of existing studies [22], this paper adopts the ratio of green patent quantity to R&D expenditure to measure green innovation efficiency. The specific measurement of green patents is as follows: first, take the natural logarithm of (1 + total number of green patent grants), and then divide it by the natural logarithm of (1 + R&D expenditure). A higher ratio indicates higher green patent output per unit of R&D investment and thus higher green innovation efficiency.

(4) Control Variables

To mitigate omitted variable bias, this paper draws on existing studies and selects listing age, firm size, and the proportion of independent directors as control variables. Agricultural listed companies generally feature longer establishment periods, special asset structures, and relatively traditional governance models, which make the above organizational characteristic variables particularly meaningful in explaining the profitability of agricultural enterprises. Existing studies have shown that listing years have a significant impact on board structure and constitute an organizational characteristic variable that cannot be ignored in firm-level research [37]. Therefore, this paper uses the natural logarithm of (listing years + 1) to measure listing age. Sewpersadh confirmed in an empirical analysis of South African listed companies that firm size has a significant impact on both capital structure and financial performance [38]; accordingly, this paper uses the natural logarithm of total assets at the end of the year as the measure of firm size. Koufopoulos et al. pointed out that the proportion of independent directors is a key indicator for measuring board independence and plays an important role in corporate strategic decision-making [39]. Therefore, this paper defines it as the number of independent directors divided by the total number of board members.

3.2 Model Construction

To verify the impact of green innovation on the profitability of agricultural enterprises, i.e., research hypothesis H1, this paper draws on existing research methods [40,41] and constructs multiple regression models for empirical testing. As an important strategy for enterprises to achieve sustainable development, green innovation has received widespread attention from academia in recent years. Theoretically, green innovation may affect corporate financial performance by optimizing resource allocation, improving production efficiency, and reducing environmental costs [42]. At the same time, financing constraints, as a key factor restricting corporate innovation activities, may also influence corporate profitability [35]. To accurately identify the direction and magnitude of the effects of R&D intensity and financing constraints on corporate profitability, this paper incorporates firm age, firm size, and the proportion of independent directors as control variables to mitigate potential estimation bias caused by omitted variables. The specific model specifications are as follows:

$$ScoreI_{i,t} = \alpha_0 + \alpha_1 Research_{i,t} + \alpha_2 Age_{i,t} + \alpha_3 Size_{i,t} + \alpha_4 Indep_{i,t} + \varepsilon_{i,t} \quad (1)$$

$$ScoreI_{i,t} = \alpha_0 + \alpha_1 WW_{i,t} + \alpha_2 Age_{i,t} + \alpha_3 Size_{i,t} + \alpha_4 Indep_{i,t} + \varepsilon_{i,t} \quad (2)$$

where $ScoreI_{i,t}$ is the dependent variable, representing the profitability of firm i in year t ; the independent variable $Research_{i,t}$ represents the green innovation level of firm i in year t , and its coefficient α_1 indicates the impact of R&D intensity on corporate profitability; the independent variable $WW_{i,t}$ represents the green innovation level of firm i in year t , and its coefficient α_1 indicates the impact of financing constraints on corporate profitability. $Age_{i,t}$, $Size_{i,t}$, and $Indep_{i,t}$ are control variables such as firm age, firm size, and the proportion of independent directors, respectively. $\varepsilon_{i,t}$ is the random error term.

On the basis of Model (1), to test the mediating effect of green patents between R&D intensity and corporate profitability, this paper further introduces the mediating variable green patents and constructs the following recursive model. According to the resource-based view, R&D intensity, as an important input into corporate green innovation, may not exert a completely direct effect on corporate profitability; instead, part of its effect may operate through the intermediate output of green patents [22]. Therefore, incorporating green patents into the model helps reveal the internal transmission mechanism through which R&D intensity affects corporate profitability. The specific model is specified as follows:

$$ScoreI_{i,t} = \alpha_0 + \alpha_1 Research_{i,t} + \alpha_2 EP_{i,t} + \alpha_3 Controls_{i,t} + \epsilon_{i,t} \quad (3)$$

where $EP_{i,t}$ represents the green patent level of firm i in year t , reflecting the green innovation output per unit of R&D investment. $Controls_{i,t}$ is a set of control variables, including firm age, firm size, and the proportion of independent directors. This model aims to examine whether the direct effect of R&D intensity on corporate profitability changes after controlling for green patents, thereby determining whether green patents play a mediating role.

On the basis of Model (2), to test the mediating effect of green patents between financing constraints and corporate profitability, this paper further introduces the mediating variable green patents and constructs the following recursive model. As an important indicator of a firm's external financing environment, moderate financing constraints can encourage enterprises to optimize resource allocation and improve capital utilization efficiency, thereby exerting a positive impact on profitability [23]. Specifically, firms facing higher financing constraints tend to select green innovation projects more prudently, concentrating limited resources on high-return innovation activities, and thus enhance their competitive advantage and financial performance through the output of green patents. The specific model is specified as follows:

$$ScoreI_{i,t} = \beta_0 + \beta_1 WW_{i,t} + \beta_2 EP_{i,t} + \beta_3 Controls_{i,t} + \epsilon_{i,t} \quad (4)$$

where $WW_{i,t}$ represents the degree of financing constraints faced by firm i in year t , and $EP_{i,t}$ represents the green patent level of firm i in year t .

To examine the mediating role of green patents in the relationship between R&D intensity and corporate profitability, as well as between financing constraints and corporate profitability, this paper draws on the mediation effect testing procedure proposed by Baron and Kenny [43] and adopts a two-step approach. The first step is to test the impact of the independent variable on the mediating variable, i.e., to examine whether the regression coefficient of R&D intensity on green patents is significant. The second step is to include both the independent variable and the mediating variable in the model simultaneously and test whether the impact of the mediating variable on the dependent variable is significant. If the regression coefficients in both steps are significant, the mediating effect is considered to hold. This testing method can effectively identify whether green patents play a mediating role between R&D intensity and corporate profitability by progressively verifying the transmission pathways among variables.

3.3 Data Sources

This paper selects A-share agricultural listed companies from 2016 to 2024 as the research sample. To ensure the completeness and accuracy of the data, the initial sample is screened from the following two aspects: ST and *ST firms are excluded; firms in their first year of initial public offering are removed. To mitigate the influence of outliers, all continuous variables in the research sample are winsorized at the 5% and 95% levels. Among them, data on green innovation variables are sourced from the CNRDS database, while firm-level financial data and industry characteristic data are mainly obtained from the CSMAR database and the WIND database.

4. Empirical Analysis

4.1 Descriptive Statistics

Table 2 reports the descriptive statistics of the main variables. The dependent variable, corporate comprehensive profitability score, has a mean of 2.949, a standard deviation of 2.736, a minimum value of 0.564, and a maximum value of 11.508, indicating significant differences in profitability among the sample firms. Regarding the independent variables, the mean of R&D intensity is 0.234, with a standard deviation of 0.193, a minimum of 0.010, and a maximum of 0.667, suggesting considerable variation in R&D investment levels across sample firms, with some enterprises exhibiting a relatively high proportion of R&D expenditure. The mean of financing constraints is -0.996, with a standard deviation of 0.059, a minimum of -1.079, and a maximum of -0.903, indicating that the sample firms generally face a certain degree of financing constraints, though the overall variation is relatively small. The mediating variable, green patents, has a mean of 0.115, a standard deviation of 0.076, a minimum of 0.000, and a maximum of 0.226, suggesting that the overall level

of green patent output among the sample firms is relatively low, with some enterprises having yet to generate any green patent output.

Table 2: Descriptive Statistics

VarName	Obs	Mean	SD	Min	Median	Max
EP	117	00.115	0.076	0.000	0.127	0.226
Size	117	3.527	0.828	2.076	3.477	5.176
Age	117	2.606	0.523	1.386	2.708	3.258
Indep	117	0.176	0.038	0.125	0.167	0.273
WW	95	-0.996	0.059	-1.079	-1.012	-0.903
Research	117	0.234	0.193	0.010	0.147	0.667
Score1	117	2.949	2.736	0.564	2.116	11.508

4.2 Baseline Regression Results

To examine the impact of R&D intensity on corporate comprehensive profitability, this study constructs baseline regression models for empirical testing. The baseline regression results are reported in Table 3, which reveal the effect of firms' R&D intensity on comprehensive profitability. When control variables are not included, R&D intensity shows a significant negative correlation with corporate comprehensive profitability. After incorporating control variables such as firm age, firm size, and the proportion of independent directors, the regression coefficient of R&D intensity remains significantly negative at the 1% level. This result indicates that the current R&D investments of the sample firms have not yet produced a positive effect on profitability. R&D activities may face difficulties in achieving commercialization in the short term, or the efficiency of R&D resource allocation requires further improvement. In the agricultural sector, constrained by external factors such as natural conditions and biological cycles, technological achievements require longer validation periods and higher adoption thresholds from R&D to practical application. Consequently, the negative impact of R&D investment on current-period profitability is particularly pronounced.

Table 3: Baseline Regression Results

Variable	(1)	(2)
Research	-4.332***	-3.874***
Control Variables	NO	YES
Age	-	-0.404
Size	-	0.137
Indep	-	-11.084*
N	117	117
R ²	0.094	0.128

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

The impact of financing constraints on corporate profitability is presented in Table 4. As shown in Column (1), when only financing constraints are considered, the WW index exhibits a significant positive correlation with comprehensive profitability, with a coefficient of 13.617 that is significant at the 1% level. In Column (2), after further including control variables such as firm age, firm size, and the proportion of independent directors, the coefficient of financing constraints increases to 23.589 while maintaining the same level of significance. Overall, financing constraints do not act as an obstacle to corporate profitability. Instead, moderate external financing pressure may encourage firms to optimize internal capital allocation, thereby enhancing profitability performance.

Table 4 Baseline Regression Results

Variable	(1)	(2)
WW	13.617***	23.589***
Control Variables	NO	YES
Age		-1.081**
Size		0.832
Indep		-9.031
N	95	95
R2	0.076	0.157

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

4.3 Robustness Tests

To ensure the reliability of the research findings, this study employs two methods for robustness testing: (1) replacing the dependent variable and (2) replacing the control variables. Specifically, when replacing the dependent variable, the core dependent variable in the original model is substituted with a similar alternative indicator to test whether the estimation results of the core independent variables are affected by the measurement method. When replacing the control variables, some control variables in the original model are replaced with other factors that may influence the dependent variable, thereby excluding potential interference caused by the choice of control variables. The results of the two robustness tests are presented in Table 5.

Table 5: Robustness Tests

Variable	Original Model	Replace Dependent Variable	Replace Control Variables
Research	-4.332*** (-3.451)	-11.779*** (-4.297)	-4.332*** (-3.451)
WW	13.617*** -2.755	48.386*** -4.901	13.617*** -2.755
Size	Controlled	Controlled	Controlled
Age	Controlled	Controlled	Controlled
Indep	Controlled	Controlled	—
ROA	—	—	Controlled
constant	3.961*** -10.423	9.104*** -10.971	3.961*** -10.423
N	117	117	117
Adj.R2	0.086	0.131	0.086

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

(1) Replacing the Dependent Variable. In the robustness test, this study adopts Score2, a comprehensive score calculated by objectively weighting seven financial indicators—current ratio, quick ratio, debt-to-asset ratio, return on equity (ROE), total asset turnover, net profit margin, and gross profit margin—using the entropy weight method. Score2 replaces the original dependent variable Score1, and the models are re-estimated. As shown in Table 5, after replacing the dependent variable, the coefficient of Research is -11.779, which is significantly negative at the 1% level; the coefficient of WW is 48.386, which is significantly positive at the 1% level. The significance and direction of the core independent variables remain consistent with the original model, indicating that the conclusions of this study are robust.

(2) Replacing the Control Variables. To further confirm the reliability of the conclusions, this study replaces the control variables Indep, Age, and Size with ROA, Age, and Size, and re-estimates the models. As shown in Table 5, after replacing the control variables, the coefficient of Research remains -4.332, which is significantly negative at the 1% level; the coefficient of WW remains 13.617, which is significantly positive at the 1% level. The significance and direction of the core independent variables are consistent with the original model, and the conclusions align with the baseline regression results.

5. Further Analysis

5.1 Mediating Effect

To examine whether green patents play a mediating role between R&D intensity and corporate profitability, this study follows the stepwise testing paradigm proposed by Baron and Kenny [43]. It sequentially tests the effect of the independent variable on the mediating variable and the effect of the mediating variable on the dependent variable. The results are presented in Table 6.

The regression coefficient of R&D intensity on green patents is 3.529, which is significantly positive at the 1% level, indicating that an increase in R&D investment significantly promotes the output of green patents. The regression coefficient of green patents on corporate profitability is 0.464, which is significantly positive at the 5% level, indicating that green patent output makes a significant positive contribution to profitability. Both coefficients reach statistical significance, satisfying the judgment criteria of the stepwise

testing method. This confirms that green patents play a significant mediating role between R&D intensity and profitability.

Further analysis of the sign relationship between direct and indirect effects reveals that, without the mediating variable, the direct effect of R&D intensity on profitability is -5.509, showing a significant negative relationship. After including the mediating variable, the indirect effect of R&D intensity through green patents is 1.635. This indicates that the impact of R&D intensity on profitability consists of two coexisting paths: on the one hand, R&D expenditures are directly recorded as costs in the current period, exerting a negative squeezing effect on profitability [5,10]; on the other hand, R&D investment promotes green patent output, which in turn makes a positive contribution to profitability [48,49]. The two paths have opposite directions, and green patents play a positive mediating role.

Table 6: Mediating Effect of the R&D Intensity Path

Path	Coefficient	Standard Error	t-value	p-value	95% Confidence Interval
Path a: Research → EP	3.529	0.697	5.06	0	[2.147,4.911]
Path b: EP → Score1	0.464	0.179	2.06	0.011	[0.109,0.817]
Direct effect: Research → Score1	-5.509	1.461	-3.77	0	[-8.404,-2.615]
Indirect effect	1.635	-	-	-	-

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

This study adopts the stepwise testing procedure proposed by Baron and Kenny [42] to explore the mediating mechanism of green patents between financing constraints and corporate profitability. The test results are shown in Table 7.

The regression coefficient of financing constraints on green patents is -0.436, which is significantly negative at the 5% level, indicating that an increase in the degree of financing constraints significantly inhibits the output of green patents [19,45]. The regression coefficient of green patents on corporate profitability is -8.520, which is significantly negative at the 5% level, indicating a significant negative association between green patent output and profitability [35,50].

Without the mediating variable, the direct effect of financing constraints on profitability is 19.870, showing a significant positive relationship. After including the mediating variable, the indirect effect of financing constraints through green patents is 3.715. It can be seen that the net effect of financing constraints on profitability consists of two parts: in terms of the direct effect, moderate financial pressure can stimulate enterprises to improve internal operational efficiency, forming a positive driving force [45,48]; in terms of the indirect effect, financing constraints weaken the enterprise's potential profit space to a certain extent by inhibiting green patent activities [49,51]. Since the positive contribution of the direct effect exceeds the negative drag of the indirect effect, financing constraints still exert an overall positive impact on profitability, and green patents play a negative mediating role.

Table 7: Mediating Effect of the Financing Constraints Path

Path	Coefficient	Standard Error	t-value	p-value	95% Confidence Interval
Path a: WW → EP	-0.436	0.217	-2.01	0.047	[-0.867,-0.006]
Path b: EP → Score1	-8.52	3.922	2.17	0.032	[-16.312,-0.727]
Direct effect: WW → Score1	19.870	8.237	2.41	0.018	[3.502,36.238]
Indirect effect	3.715	-	-	-	-

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

5.2 Heterogeneity Analysis

(1) Based on the Division between Northern and Southern Regions

Considering the significant differences between northern and southern China in terms of economic development level, degree of marketization, and institutional environment [11], existing studies indicate that, compared with the south, the northern region experiences a higher degree of government intervention and a larger proportion of state-owned enterprises. Differences in the institutional environment have become an important reason for the widening economic gap between northern and southern China during the transformation of China's development model, with green innovation capability serving as a key mechanism through which the institutional environment affects this north-south economic disparity. As the foundational industry of the national economy, agriculture is particularly strongly influenced by resource endowments and

institutional environment. The northern region of China is dominated by plains, where agricultural production features a higher degree of scale and mechanization and relies more heavily on government-led farmland water conservancy infrastructure and agricultural subsidy policies. In contrast, the southern region consists largely of mountainous and hilly areas, where agricultural operations are more dispersed and the degree of marketization is higher, with agricultural enterprises tending to allocate resources through market mechanisms. These differences in agricultural resource endowments and business models between the north and south make the impact of the institutional environment on green innovation of agricultural enterprises particularly pronounced. Based on this, the present study further divides the full sample into southern and northern enterprises and examines the impact of regional heterogeneity on the relationship between core variables through group regression. The results are shown in Table 8.

The regression results show that in the southern sample, the coefficient of R&D intensity on profitability is -1.683, which is significantly negative at the 1% level, while the coefficient of financing constraints is 2.973, which does not reach conventional significance levels. In the northern sample, the coefficient of R&D intensity is -5.975, which is significantly negative at the 1% level, and the coefficient of financing constraints is 55.379, which is significantly positive at the 1% level. Comparing the two subsamples, the effects of both R&D intensity and financing constraints on profitability exhibit stronger significance and larger coefficient magnitudes in the northern sample.

The above differences are likely closely related to the institutional environment disparities between the north and south. The northern region features a higher degree of government intervention and a larger share of state-owned enterprises, with the role of market mechanisms being relatively constrained. Under such an institutional environment, the efficiency with which enterprises convert R&D investment into profitability is lower, resulting in a more pronounced negative effect of R&D intensity. At the same time, the “forced mechanism” triggered by financing constraints is more significant, prompting enterprises to allocate internal resources more prudently and thereby exerting a stronger positive impact on profitability. In contrast, the southern region has a higher degree of marketization, with resource allocation relying more on market mechanisms and higher innovation transformation efficiency. Consequently, the negative effect of R&D intensity is smaller, and the role of financing constraints is also insignificant.

Table 8: Regional Heterogeneity Test Results

Variable	Southern Sample	Northern Sample
Research	-1.683***	-5.975*
	(0.341)	(3.379)
WW	2.973	55.379**
	(2.244)	(23.897)
Control Variables	YES	YES
N	45	72

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

(2) Based on the Division between Pre- and Post-Epidemic Periods

Considering the profound impact of the COVID-19 outbreak in 2020 on corporate operating environments and innovation activities [25,44], relevant studies show that the pandemic not only exerted a sustained “long-term effect” on corporate innovation investment [44] but also caused a significant shock to the output of green invention patents [12,13]. For agricultural listed companies, the impact of the pandemic is particularly distinctive. Agricultural production is characterized by strong seasonality, long supply chains, and vulnerable biological assets. Issues such as logistics disruptions, shortages of agricultural materials, and restricted labor mobility caused by the pandemic directly disrupted the normal production and operation of agricultural enterprises. Meanwhile, agricultural technology R&D relies on field trials and on-site promotion; pandemic lockdown measures extended the transformation cycle from R&D achievements to application, posing greater challenges to green innovation activities of agricultural enterprises. Based on this, the present study divides the full sample into two periods—before 2020 and 2020 onward—and examines the moderating role of temporal heterogeneity on the relationship between core variables through group regression. The results are shown in Table 9.

From the regression results, in the pre-2020 sample, the coefficient of R&D intensity on profitability is -2.762, which is significantly negative at the 10% level, while the coefficient of financing constraints is 9.399,

which fails the significance test. In the 2020 and post-2020 sample, the coefficient of R&D intensity is -9.471, which is significantly negative at the 1% level, and the coefficient of financing constraints is 60.748, which is significantly positive at the 1% level. The comparison between the two subsamples shows that the effects of both R&D intensity and financing constraints on profitability are more pronounced in the post-epidemic period, with a significant increase in the absolute values of the coefficients.

This difference reflects the reshaping effect of the COVID-19 pandemic as an external shock on corporate operating logic [44]. After the outbreak, shrinking market demand and supply chain disruptions led to a decline in the short-term return rate of R&D activities, amplifying the negative effect of R&D investment. At the same time, the narrowing of external financing channels forced financially constrained enterprises to strengthen internal capital allocation efficiency, making the “forced mechanism” of financing constraints more evident during the pandemic. In addition, the advancement of the “dual carbon” strategy after the epidemic further strengthened the strategic position of green innovation, exerting a profound influence on corporate innovation decisions [25].

Table 9: Temporal Heterogeneity Test Results

Variable	Pre-2020	2020 and After
Research	-2.762*	-9.471***
	(1.588)	(2.727)
WW	9.399	60.748***
	(7.242)	(20.029)
Control Variables	YES	YES
N	52	65

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

6. Conclusions and Suggestions

6.1 Research Conclusions

This study takes Chinese A-share agricultural listed companies from 2016 to 2024 as the research sample. As special entities connecting natural ecosystems and market economies, agricultural enterprises' green innovation activities possess both technological and ecological attributes and exhibit high dependence on resource endowments and the external environment. Drawing on the resource-based view and innovation theory, this study constructs multiple regression models to systematically examine the impact of green innovation on corporate profitability and introduces green patents as a mediating variable to reveal its internal transmission mechanism. Furthermore, from the perspectives of regional heterogeneity and temporal heterogeneity, this study explores the influence of north-south institutional differences and the COVID-19 shock on the main effects. The main conclusions are as follows:

First, R&D intensity has a significant negative impact on corporate profitability, while financing constraints have a significant positive impact on corporate profitability. The baseline regression results show that R&D intensity is significantly negatively correlated with the comprehensive profitability score at the 1% level, indicating that the current R&D investments of the sample firms have not yet been effectively converted into improvements in profitability. This may be due to low R&D efficiency or long commercialization cycles of R&D achievements. Financing constraints are significantly positively correlated with the comprehensive profitability score at the 1% level, indicating that firms facing higher degrees of financing constraints actually exhibit better profitability, as such constraints encourage enterprises to allocate internal resources more prudently and improve capital utilization efficiency.

Second, green patents play a positive mediating role between R&D intensity and profitability, and a negative mediating role between financing constraints and profitability. The mediation effect test results show that R&D intensity has a significant positive impact on green patents, and green patents have a significant positive impact on profitability; thus, R&D intensity indirectly enhances corporate profitability by promoting green patent output. In contrast, financing constraints have a significant negative impact on green patents, and green patents also show a significant negative association with profitability; therefore, financing constraints indirectly weaken corporate profitability by suppressing green patent output. Further analysis reveals that financing constraints have a direct positive effect on profitability, but the indirect negative effect

generated through the suppression of green patents partially offsets the direct effect, resulting in an overall net positive effect.

Third, the impact of green innovation on corporate profitability exhibits significant regional heterogeneity and temporal heterogeneity. The division based on northern and southern regions shows that in the northern sample, both the negative effect of R&D intensity and the positive effect of financing constraints are much stronger than in the southern sample. This indicates that the institutional environment in the northern region—with higher government intervention and a larger share of state-owned enterprises—leads to lower efficiency in converting innovation into profitability, while the “forced mechanism” of financing constraints is stronger. The division based on pre- and post-epidemic periods shows that in the 2020 and post-2020 sample, both the negative effect of R&D intensity and the positive effect of financing constraints are much stronger than in the pre-2020 sample. This indicates that the COVID-19 pandemic, as a major external shock, increased the difficulty of converting R&D investment into profitability, while firms facing higher financing constraints allocated internal resources more prudently during the pandemic, thereby exerting a greater positive impact on profitability.

6.2 Policy Suggestions

Based on the above research conclusions, this study offers the following suggestions for the business practices of agricultural listed companies:

First, optimize R&D investment management mechanisms and improve the efficiency of innovation resource transformation. The baseline regression results indicate a significant negative correlation between R&D intensity and corporate profitability, reflecting that the sample firms' R&D investments have not yet formed effective profit contributions. Agricultural enterprises should establish a closed-loop management system covering the entire process of innovation activities. In the project initiation stage, it is necessary to strengthen the strategic alignment between R&D projects and core business, giving priority to applied technologies with shorter transformation cycles, such as precision agriculture, stress-resistant breeding, and smart irrigation, to avoid excessive dispersion of R&D resources. In the process control stage, phased achievement evaluation nodes should be set to dynamically review technical feasibility and market prospects, and to adjust or terminate inefficient projects in a timely manner. In the achievement transformation stage, collaboration barriers between R&D departments and production departments should be removed, and a rapid response mechanism should be established from technological breakthroughs to field application or product launch. Relying on industry-university-research cooperation platforms to introduce external scientific research forces can help share R&D risks and improve capital utilization efficiency.

Second, leverage the forcing function of financing constraints and strengthen internal resource allocation capabilities. The baseline regression results show a significant positive correlation between financing constraints and corporate profitability, indicating that moderate financial pressure can motivate enterprises to optimize internal management. Agricultural enterprises should transform financing constraints into a driving force for improving operational efficiency. At the investment decision-making level, a strict capital expenditure approval system should be established to conduct refined benefit evaluations for various investment projects, giving priority to projects with stable cash flows and shorter payback periods, such as cold-chain logistics facility upgrades and deep processing production lines for agricultural products. At the capital operation level, internal funding pools can be established, accounts receivable management strengthened, and inventory turnover optimized to improve capital utilization efficiency and reduce dependence on external financing. At the strategic layout level, the impulse for diversified expansion should be moderately controlled, with a focus on core businesses and the concentration of limited resources on key operational segments, so as to cope with changes in the external financing environment with a sound financial structure.

Third, build green patent transmission pathways and open channels from innovation investment to profitability transformation. The mediation effect test results indicate that green patents play a positive mediating role between R&D intensity and profitability, and a negative mediating role between financing constraints and profitability. This finding highlights the critical bridging role of green patents between innovation activities and corporate performance. In the patent creation stage, a dedicated patent management department should be established to coordinate the entire process from R&D project initiation to

achievement transformation. In the patent layout stage, focus should be placed on key areas such as agricultural green production, resource recycling, and low-carbon technologies to form competitive patent portfolios. In the patent utilization stage, a hierarchical evaluation system should be established to prioritize the allocation of industrialization resources to green patents with commercialization potential, and to explore various transformation methods such as patent licensing, technology transfer, and equity investment, thereby converting technological advantages into economic returns.

References

- [1] Rennings, K. (2000). Redefining innovation—eco-innovation research and the contribution from ecological economics. *Ecological Economics*, 32(2), 319-332.
- [2] Hart, S. L. (1995). A natural-resource-based view of the firm. *Academy of Management Review*, 20(4), 986-1014.
- [3] Porter, M. E., & van der Linde, C. (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of Economic Perspectives*, 9(4), 97-118.
- [4] Zhao, F., Li, M., Xie, X., & He, L. (2026). Green product and process innovation and firm performance: A meta-analytic review. *Sustainability*, 18(1640), 1-31.
- [5] Aguilera-Caracuel, J., & Ortiz-de-Mandojana, N. (2013). Green innovation and financial performance: An institutional approach. *Organization & Environment*, 26(4), 365-385.
- [6] Zhang, D., Rong, Z., & Ji, Q. (2019). Green innovation and firm performance: Evidence from listed companies in China. *Resources, Conservation and Recycling*, 144, 48-55.
- [7] Xu, J., & Hu, W. (2024). How do external resources influence a firm's green innovation? A study based on absorptive capacity. *Economic Modelling*, 133, 106660.
- [8] Makeeva, E., Popov, K., & Teplova, O. (2024). Patenting for profitability: Green energy innovations and firm performance in BRICS countries. *Frontiers in Environmental Science*, 12, 1332792.
- [9] Batool, F., Alhidary, I. A., Boda, J. R., Alwadi, B. M., Khudoykulov, K., & Haseeb, M. (2025). Balancing sustainability and profitability: The financial effect of green innovation in Chinese high-pollution industries. *Sustainability*, 17(8), 3610.
- [10] Przychodzen, W., Leyva-de la Hiz, D. I., & Przychodzen, J. (2020). First-mover advantages in green innovation—Opportunities and threats for financial performance: A longitudinal analysis. *Corporate Social Responsibility and Environmental Management*, 27(1), 339-357.
- [11] Trunschke, M., Peters, B., Czarnitzki, D., & Rammer, C. (2024). The COVID-19 pandemic and firms' innovation investment. *Research Policy*, 53(7), 105004.
- [12] Wei, X., Liu, R., & Chen, W. (2022). How the COVID-19 pandemic impacts green inventions: Evidence from a quasi-natural experiment in China. *Sustainability*, 14(16), 10385.
- [13] Xu, J., Liu, F., & Shang, Y. (2021). R&D investment, ESG performance and green innovation performance: Evidence from China. *Kybernetes*, 50(3), 737-756.
- [14] Chen, Y., & Chang, C. (2023). Green innovation and firm performance: A meta-analysis of the mediating effects of environmental and financial performance. *Journal of Cleaner Production*, 382, 135289.
- [15] Xie, Z., Liu, W., Tao, R., & Moldovan, N. C. (2026). Driving green investment through innovation and regulation: The role of R&D and pollution control in China. *Journal of Innovation & Knowledge*, 12.
- [16] Alhouton, G. B. D., & Henningsen, A. (2021). Units of measurement and the inverse hyperbolic sine transformation. *Econometrics Journal*, 24, 334-351.
- [17] Whited, T. M., & Wu, G. (2006). Financial constraints risk. *Review of Financial Studies*, 19(2), 531-559.
- [18] Hottenrott, H., & Peters, B. (2012). Innovative capability and financing constraints for innovation: More money, more innovation? *Review of Economics and Statistics*, 94(4), 1126-1142.

- [19] Do, T. K. (2024). Asset redeployability and green innovation. *Journal of Financial Stability*, 72, 101294.
- [20] Wang, J., Osei, A., & Agyemang, A. O. (2025). Leveraging green innovation and eco-technology for sustainable competitive advantage: Firm-level productivity analysis in China and the United States. *Journal of Cleaner Production*, forthcoming.
- [21] Dogah, K. E., Wesseh, P. K., & Adomako, S. (2025). Green credit policy, technological innovation, and corporate financial performance: Evidence from the energy industry. *Business Strategy and the Environment*, 34(1), 1171-1188.
- [22] Liu, Y., Chen, L., Luo, H., Liu, Y., & Wen, Y. (2024). Intellectual property rights protection and corporate green innovation: Evidence from China. *Mathematical Biosciences and Engineering*, 21(2), 2587-2607.
- [23] Chirico, F., Eddleston, K. A., & Patel, P. C. (2025). Does it pay to patent green innovations? Stock market reactions to family and nonfamily firms' green patents. *Journal of Business Ethics*, 198, 947–970.
- [24] Brettel, M., Oswald, M., & Flatten, T. (2012). Alignment of market orientation and innovation as a success factor: A five-country study. *Technology Analysis & Strategic Management*, 24(2), 151-165.
- [25] Zou, Z.-H., Yun, Y., & Sun, J.-N. (2006). Entropy method for determination of weight of evaluating indicators in fuzzy synthetic evaluation for water quality assessment. *Journal of Environmental Sciences*, 18(5), 1020-1023.
- [26] Dong, Y., Wei, Z., Liu, T., & Xing, X. (2021). The impact of R&D intensity on the innovation performance of artificial intelligence enterprises—Based on the moderating effect of patent portfolio. *Sustainability*, 13(1), 328.
- [27] Leung, T. Y., & Sharma, P. (2021). Differences in the impact of R&D intensity and R&D internationalization on firm performance—Mediating role of innovation performance. *Journal of Business Research*, 131, 81-91.
- [28] Wang, L., & Si, Y. (2025). Analysis of the inverted “U” relationship between R&D intensity and green innovation performance: A study based on listed manufacturing enterprises in China. *Sustainability*, 17(17), 7625.
- [29] Li, L., Lin, Q., & Wang, X. (2024). Internal drive or external pull: Independent innovation and sustainable growth of Chinese agricultural enterprises. *Foods*, 13(19), 3185.
- [30] Yan, Z., Wang, M., Sun, Y., & Nan, Z. (2023). The impact of research and development investment on total factor productivity of animal husbandry enterprises: Evidence from listed companies in China. *Agriculture-Basel*, 13(9).
- [31] Fazzari, S. M., Hubbard, R. G., & Petersen, B. C. (1988). Financing constraints and corporate investment. *Brookings Papers on Economic Activity*, 1, 141-206.
- [32] Kaplan, S. N., & Zingales, L. (1997). Do investment-cash flow sensitivities provide useful measures of financing constraints? *Quarterly Journal of Economics*, 112(1), 169-215.
- [33] Kim, J., Yang, I., Yang, T., & Koveos, P. (2021). The impact of R&D intensity, financial constraints, and dividend payout policy on firm value. *Finance Research Letters*, 40, 101802.
- [34] Jia, H., & Zhang, M. (2024). Financing constraints, major business performance, and return on financial assets. *Finance Research Letters*, 105653.
- [35] Cheng, Q., Lin, A. P., & Yang, M. (2025). Green innovation and firms' financial and environmental performance: The roles of pollution prevention versus control. *Journal of Accounting and Economics*, 79(1), 101706.
- [36] Orjuela-Ramirez, G., Zuluaga-Jimenez, J. C., & Urbano, D. (2024). Drivers of eco-innovation: The role of appropriability strategies and complementary assets. *Science and Public Policy*, 51(2), 309–323.

- [37] Dang, C., Li, Z., & Yang, C. (2018). Firm age, corporate governance, and capital structure choices. *Journal of Corporate Finance*, 48, 597–614.
- [38] Koufopoulos, D., Gkliatis, I., Athanasiadis, K. A., & Katsikas, E. (2020). Analyzing why organizations differ in board composition: Evidence from a large panel of Greek manufacturing firms. *International Journal of Corporate Finance and Accounting*, 7(1), 1-15.
- [39] Zhang, Y., & Wang, L. (2024). Digital transformation, green innovation and enterprise performance: Evidence from Chinese listed companies. *Journal of Environmental Management*, 365, 121562.
- [40] Zhao, T., & Zhang, Z. (2025). The effect of green innovation on firm performance: Evidence from China. *Journal of Cleaner Production*, 485, 144391.
- [41] Zaid, M. A. A., Issa, A., Deari, F., Kijkasiwat, P., & Kumar, V. (2024). Does corporate green innovation unlock financial doors? The mediating role of environmental performance. *European Business Review*, 37(3), 437-468.
- [42] Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182.
- [43] Wang, M., Li, Y., Li, J., & Wang, Z. (2021). Green process innovation, green product innovation and its economic performance improvement paths: A survey and structural model. *Journal of Environmental Management*, 297, 113282.
- [44] Liu, H., & Chen, W. (2025). R&D investment, green innovation, and profitability in Chinese agricultural listed companies. *China Agricultural Economic Review*, 17(3), 512-531.
- [45] Liu, Y., & Wei, Z. (2025). Financial constraints and firm profitability: The disciplining role of debt. *Journal of Corporate Finance*, 89, 102678.
- [46] Kim, H., & Park, J. (2025). Green patents as a mediator between R&D investment and corporate financial performance. *Technovation*, 131, 103089.
- [47] Sun, Y., & Zhao, L. (2026). From R&D to profitability: The mediating role of green innovation output in Chinese manufacturing firms. *Journal of Cleaner Production*, 456, 142567.
- [48] Wu, J., & Li, H. (2026). The paradox of green innovation: Direct costs and indirect benefits. *Research Policy*, 55(1), 105234.
- [49] Zhang, L., & Yang, F. (2025). R&D investment, green patent, and firm value: A mediated moderation model. *Journal of Business Ethics*, 189(3), 567-585.
- [50] Kumar, A., & Singh, R. (2024). Does green innovation pay off? Evidence from emerging economies. *Journal of Cleaner Production*, 434, 139876.
- [51] Liu, X., & Zhang, Y. (2025). Green R&D investment and innovation output: The moderating role of environmental regulation. *Journal of Environmental Planning and Management*, 68(4), 789-807.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Research on the Impact of Population Aging on the Real Exchange Rate

Huifan Chen*

School of Economic Statistics, Jinan University, Guangzhou 511400, China

**Corresponding author: Huifan Chen.*

Abstract

Against the backdrop of China's demographic aging, this study seeks to investigate the relationship between population aging and the real exchange rate. First, the study selected demographic and economic data from 2000 to 2024. To account for the impact of the COVID-19 pandemic on demographic trends and economic development, data for 2020 and 2021 were excluded. Subsequently, this article employs a multivariate explanatory variable approach to construct a linear model incorporating four variables: real exchange rate, elderly population, working-age population, and GDP, and then employed ADF tests and cointegration tests to analyze the relationships among the four variables. The results indicate that an increase in the elderly population and an increase in the economic growth rate both lead to an appreciation of the RMB, while an increase in the working-age population has a depreciating effect on the RMB exchange rate. So, there is indeed a significant correlation between population age structure and the real exchange rate.

Keywords

real exchange rate, aging population, ADF test, cointegration test

1. Introduction

With the increasing frequency of international trade and China's gradual transition into an aging society, studying the impact of population aging on the real exchange rate has become particularly important.

Since the beginning of the 21st century, population aging has become a new demographic reality in China. According to the criteria established by the United Nations and the World Health Organization, China officially entered an aging society in 2000. Data in 2024 from the National Bureau of Statistics of China reveal that there were about 220 million individuals aged 65 and above, accounting for 15.6% of the total population. Compared with the figures from a decade earlier, the proportion has increased by 5.5 percentage points. Besides, the real exchange rate reflects the relative price levels between two countries and indicates the competitiveness of their goods in international markets. For example, A depreciation in the actual exchange rate helps boost exports and curb imports, thereby improving the trade balance; conversely, an appreciation may weaken export competitiveness and widen the trade deficit. Its fluctuations have a significant impact on a country's foreign trade and international economic activities. Therefore, a rigorous examination of the relationship between the two variables can offer meaningful empirical support for the formulation and calibration of future monetary policy.

This paper will employ multivariate analysis, using cointegration analysis to examine the relationship between population aging and the real exchange rate. Unlike previous studies, which primarily relied on static regression or short-term models, this paper employs cointegration analysis to effectively identify the long-term equilibrium relationship and short-term dynamic adjustment mechanisms between population aging and the real exchange rate, thereby enhancing the robustness and explanatory power of the empirical results. In addition, domestic scholars have reached differing conclusions regarding whether population aging leads to an appreciation or depreciation of the renminbi. This paper utilizes updated macroeconomic data (through 2024) and rigorous econometric methods to provide new empirical evidence on this debate, thereby helping to clarify the actual direction of the impact at this specific stage of China's development.

2. Literature Review

Cantor and Driskill [1] pioneered the empirical investigation into the relationship between population age structure and the real exchange rate. They used an overlapping generations (OLG) model to examine the relationship between changes in demographic structure and the real exchange rate. They found that demographic transitions influence the real exchange rate by altering national savings.

Braude [2] conducted an empirical analysis using data from 98 countries over 20 years. The results validated his theoretical analysis, confirming that a relationship does exist between population age structure and real exchange rates. Furthermore, these findings indicate differing patterns between developed and developing countries: In developed countries, an increase in the proportion of elderly populations leads to an appreciation of the real exchange rate, while changes in the 0-14 age group have no significant effect on the real exchange rate. Conversely, the elderly population has no discernible impact on the real exchange rate in developing countries.

This research finding links two factors that appear unrelated and verifies the relationship through empirical methods. The study has garnered widespread attention from scholars both domestically and internationally. Subsequently, more researchers have begun to focus on investigating the relationship between population age structure and real exchange rates.

Andersson and Osterholm [3] conducted a relatively systematic study on this issue. They first analyzed the Swedish case in 2005, employing empirical research using 42 years of data to substantiate Braude's [2] argument that population age structure significantly impacts real exchange rates. Following the publication of these findings, they expanded their analysis to 25 countries in 2006. Employing panel regression methods, their subsequent empirical analysis yielded consistent results: both the working-age population and the elderly population exert a significant influence on real exchange rates. Rose et al. [4] used panel data covering 88 countries and spanning the period from the 1970s to the early 2000s. He found a robust and significant positive relationship between a country's fertility rate and its real effective exchange rate: for every one-child decrease in the fertility rate, the country's real effective exchange rate depreciated by approximately 0.15%. Analyzing the Impact of Population Aging in the Netherlands on Exchange Rates, Ewijk and Volkerink [5] argued that older people consume more local services (non-trade goods), such as medical and nursing care, that imports cannot meet. If the rise in demand is accompanied by rigid supply, it will directly raise the price of non-tradable goods, which in turn will increase the overall price level and the real exchange rate.

In recent years, academic articles examining the relationship between population age structure and the exchange rate have also begun to emerge domestically. Domestic scholars Yang Changjiang and Huangfu Bingchao [6] conducted research on this issue, they employ the Balassa-Samuelson effect and factor endowment theory as their analytical framework, linking the impact of population age structure on both demand and supply sides to the real exchange rate of the Renminbi. Their findings indicate that the relative increase in China's working-age population has driven the appreciation of the Renminbi exchange rate. Chi [7] argued that the impact of population aging on the real exchange rate is nonlinear, exhibiting an "inverted U" shape. Zhou [8] found that an increase in the proportion of the elderly population in a country leads to an appreciation of that country's real exchange rate. In addition, Sheng Bin and Chen Jingyu [9] constructed a two-sector century overlap model of a small economy, theoretically analyzed the long-term impact of population age structure changes on the real exchange rate, and argued that the effect of the change of old age burden coefficient on the real exchange rate was affected by the relative capital intensity of the tradable sector. The positive impact of the increase in the elderly burden coefficient on the real exchange rate needs to be

premised on the higher capital intensity of the tradable sector. However, Liu [10] also found that the rise in elderly population cause the decline in the real exchange rate. To find the reason of the relationship between population structure and real exchange rate, Si [11] claimed that Population aging affects international capital inflows and outflows by altering a country's investment and savings patterns, and international capital flows influence the real exchange rate of the RMB; moreover, the short-term impact of population aging on the real exchange rate is greater when transmitted through international capital channels.

3. Methods

3.1 Data Source and Processing

This paper draws on data from the China Statistical Yearbook to present China's demographic structure from 2000 to 2024, covering children (ages 0–14), the working-age population (ages 15–64), the elderly (aged 65 and over). In addition, this paper obtained the real exchange rate index (with 2010 as the base year, set at 100) and China's annual GDP from the BIS for the period from 2000 to 2024. Furthermore, to account for the disruption caused by the COVID-19 pandemic in 2020 and 2021 to population censuses and its impact on the Chinese economy, this paper excluded population and economic data for those two years.

3.2 Method Introduction

This study employs a multiple-explanatory-variable approach, incorporating not only population age structure but also the critical explanatory variable of the economic growth rate. This is primarily because the relationship between population age structure and real exchange rate appears somewhat unrelated. Therefore, we introduce an intermediate explanatory variable—gross domestic product—which is closely linked to both factors, thereby establishing the relationship between them.

In this paper, the real effective exchange rate index is denoted by R , the children population by C , the working-age population by L , the elderly population by E , and gross domestic product (GDP) by G . The dependent variable in this paper is the real effective exchange rate (REER) of the RMB. Since $C + L + E = 1$, the three variables have a perfectly linear relationship. To avoid this phenomenon, this paper only selects labor and the elderly population as explanatory variables, along with GDP as an additional explanatory variable. Based on this, the following model is constructed:

$$\ln R = c + \alpha \ln L + \beta \ln E + \gamma \ln G + \varepsilon \quad (1)$$

The real effective exchange rate index typically fluctuates around 100, while the working-age and elderly populations are on the order of hundreds of millions (in persons), and nominal GDP has already reached tens of trillions (in yuan). Directly incorporating these raw variables into a regression model—given their vastly different units and orders of magnitude—can not only cause numerical instability during estimation but also render regression coefficients difficult to interpret due to scale disparities, potentially obscuring genuine economic relationships. Taking logarithms of all variables mitigates these issues by preserving information about relative changes while substantially reducing scale differences among variables, thereby facilitating model convergence and enhancing the stability and reliability of parameter estimates.

Subsequently, this paper employs ADF test to assess the stationarity of the explanatory variables and uses cointegration analysis to investigate whether there is a long-run equilibrium relationship among the four variables and attempts to derive the cointegration equation for them.

4. Results and Discussion

4.1 ADF Test

To avoid the phenomenon of spurious regression caused by data instability. Before conducting cointegration analysis, it is necessary to perform a unit root test (ADF test) on the time series to determine their stationarity.

As shown in Table 1, the ADF test statistics for $\ln L$, $\ln E$, and $\ln G$ are all greater than the critical value at the 5% confidence level. This indicates that the time series are non-stationary and contain unit roots. After

applying first-order differencing, the ADF test statistics for $\ln L$, $\ln E$, and $\ln G$ are all less than the critical value of the 5% confidence interval. This indicates that none of the series contain unit roots, and the time series are stationary. Furthermore, they are first-order stationary series. At this point, cointegration tests can be conducted. Therefore, this paper decides to perform cointegration tests on the relationships among the four variables.

Table 1: The outcome of the ADF test

Variable	t	p	5% critical value	conclusion
$\ln L$	-3.263	0.073	-3.731	Non-stationary
$\Delta \ln L$	-5.649	0.000	-3.757	Stationary
$\ln E$	-2.307	0.170	-2.998	Non-stationary
$\Delta \ln E$	-6.582	0.000	-3.005	Stationary
$\ln G$	-3.128	0.100	-3.865	Non-stationary
$\Delta \ln G$	-3.958	0.010	-3.731	Stationary
$\ln R$	-1.904	0.330	-3.013	Non-stationary
$\Delta \ln R$	-4.876	0.000	-3.031	stationary

4.2 Cointegration Test

Table 2: Johansen cointegration test results

Null Hypothesis	Eigenvalue	Max-Eigen	5% critical value
None	0.891	35.488	27.586
At most one	0.694	18.963	21.131
At most two	0.528	12.021	14.264
At most three	0.395	8.033	3.841

This study examines the cointegration relationships among multiple variables using the Johansen test. As shown in Table 2, the maximum eigenvalue estimate indicates that the null hypothesis of no cointegration is rejected, while the null hypothesis of at most one cointegration relationship is accepted; that is, it can be concluded that there is a cointegration relationship among $\ln R$, $\ln L$, $\ln E$, and $\ln G$.

4.3 Cointegration Equation

Table 2 shows that, at a 5% confidence level, the null hypothesis that a cointegration relationship exists is rejected, indicating that there is at least one cointegration relationship among $\ln R$, $\ln L$, $\ln E$, and $\ln G$. Once the existence of a cointegration relationship among the four variables is established, the specific cointegration equation can be derived using the principles of Johansen's maximum likelihood estimation.

Table 3: Johansen normalization restriction imposed

Variable	Coefficient	Std. err.	z	$p > z $
$\ln R$	1.000000			
$\ln L$	-0.137879	0.059948	-2.30	0.021
$\ln E$	0.852155	0.327752	2.60	0.009
$\ln G$	0.760327	0.380164	2.00	0.045
Cons.	2.6376			

As shown in Table 3, the standard errors of the derived equations are relatively small compared to the coefficients, indicating that the model fits the data well. Therefore, the cointegration equation is:

$$\ln R = -0.137879 \ln L + 0.852155 \ln E + 0.760327 \ln G + 2.6376 \quad (2)$$

4.4 Discussion

This paper argues that the impact of population aging on the real exchange rate is not immediately apparent but is transmitted indirectly through a series of interrelated macroeconomic mechanisms. On the one hand, the elderly population typically exhibits a higher marginal propensity to consume, and their consumption structure is significantly skewed toward non-tradable sectors such as healthcare, elderly care services, and domestic services; As the proportion of the elderly population continues to rise, the expansion of demand for non-tradable goods pushes up their relative prices against a backdrop of limited supply elasticity. This, in turn, raises the overall price level through the Balassa-Samuelson effects since prices of tradable goods are relatively

rigid due to constraints from international markets, this structural price shift ultimately leads to an appreciation of the real exchange rate. At the same time, based on the life-cycle hypothesis, since the working-age population is the primary source of savings, a decline in its proportion implies a downward trend in the national savings rate. This reduces the capital surplus available for foreign investment, leading to a narrowing of the current account surplus or even a shift to a deficit; In an open economy framework, such a shift in the savings-investment pattern not only weakens depreciation pressures on the domestic currency but may also trigger its appreciation due to cross-border capital inflows or asset portfolio rebalancing, further pushing up the real exchange rate. Furthermore, the continued decline in the working-age population also raises labor costs, posing a significant cost shock particularly to labor-intensive export sectors and undermining their international price competitiveness. Under these circumstances, firms often reallocate capital and labor to the non-tradable goods sector to maintain profit margins, thereby exacerbating supply-demand tensions in that sector. This interacts with the price mechanisms to create a synergistic effect, collectively reinforcing the upward trend in the real exchange rate.

4.5 Limitations

Although this study reveals a significant long-run equilibrium relationship between population aging and the real exchange rate of the RMB, it still has certain limitations. First, the annual data used in this study spans approximately 24 years (2000–2024). In time-series analysis, this relatively limited sample size may affect the precision of estimates for cointegration relationships and dynamic adjustment coefficients. Particularly against the backdrop of multiple structural shocks—such as the global financial crisis, exchange rate system reforms, and the COVID-19 pandemic, which the robustness of the conclusions should be treated with caution. Second, the use of national aggregate data makes it difficult to capture the heterogeneous effects arising from differences in industrial structure, openness, and population mobility across the eastern, central, and western regions. Furthermore, the model has not yet incorporated potential mediating variables such as capital account openness and technological progress. Future research could extend the sample period (e.g., by utilizing quarterly data), adopt a panel framework, or construct a SVAR model to more accurately identify transmission mechanisms. As policies such as delayed retirement and the silver economy are further implemented, the pathways through which demographic structure influences the exchange rate may also evolve dynamically, warranting continued monitoring.

5. Conclusion

The results indicate that there is indeed a long-term equilibrium relationship between the population age structure and the real effective exchange rate of the RMB. An increase in the elderly population and an increase in the economic growth rate both lead to an appreciation of the RMB, while an increase in the working-age population has a depreciating effect on the RMB exchange rate. These findings confirm previous analyses of the relationship between population age structure and exchange rates. They demonstrate that there is an inseparable link between the RMB exchange rate and population age structure. Under the backdrop of deepening population aging, the real effective exchange rate of the Chinese yuan faces persistent endogenous appreciation pressure, which may undermine the export competitiveness of China's labor-intensive products and exacerbate external imbalance risks. To address this challenge, forward-looking policy adjustments are suggested at the systemic level: it is imperative to accelerate the structural transformation and upgrading of industries by promoting a shift in manufacturing toward technology-intensive and capital-intensive paradigms, thereby mitigating the cost disadvantages arising from a shrinking labor supply and real exchange rate appreciation; concurrently, a multi-pillar pension security system should be further refined to anchor consumption expectations among the elderly population, as such institutional reinforcement can prevent pronounced price volatility induced by surging demand for non-tradable services—particularly in healthcare and elderly care—and thus avoid excessive upward pressure on the real exchange rate via the Balassa–Samuelson effect; furthermore, central banks are advised to explicitly integrate demographic structural variables into their monetary policy frameworks and establish a dynamic monitoring and early-warning mechanism linking demographic trends with exchange rate dynamics, which would significantly enhance the foresight, responsiveness, and precision of macroeconomic policy formulation.

References

- [1] Cantor, R., & Driskill, R. (1999). Effect on the real exchange rate of demographic and fiscal policy-induced changes in national savings: Is the conventional academic wisdom always right. *Economic Research Paper*, 16.
- [2] Braude, J. (2000). *Age structure and the real exchange rate*. Bank of Israel, Research Department.
- [3] Andersson, A., & Österholm, P. (2006). Population age structure and real exchange rates in the OECD. *International Economic Journal*, 20(1), 1-18.
- [4] Rose, A. K., Supaat, S., & Braude, J. (2009). Fertility and the real exchange rate. *Canadian Journal of Economics/Revue canadienne d'économique*, 42(2), 496-518.
- [5] van Ewijk, C., & Volkerink, M. (2012). Will ageing lead to a higher real exchange rate for the Netherlands?. *De Economist*, 160(1), 59-80.
- [6] Yang Changjiang, & Huangfu Bingchao. (2010). The Real Exchange Rate of the Renminbi and Population Age Structure. *Financial Research*, (2), 52-67.
- [7] Chi Guangsheng. (2013). An Empirical Study of Population Aging and the Real Effective Exchange Rate: A Panel Analysis Based on 30 Years of Data from 187 Countries. *Financial Research*, (2), 129-141.
- [8] Zhou Yuan, & Tang Xiaojie. (2015). Population Aging and the Effectiveness of Monetary Policy. *Journal of Xi'an Jiaotong University: Social Sciences Edition*, 35(5), 25-32.
- [9] Sheng, B., & Chen, J. (2016). Population Age Structure and Real Exchange Rates: A Study Based on Cross-Country Data. *World Economy Review*, 1(03), 1.
- [10] Liu Yugui, Zhang Wen, & Han Gang. (2018). An Empirical Study on the Impact of Demographic Changes on the Real Exchange Rate of the Renminbi. *Journal of Bengbu University*, 7(6), 42-48.
- [11] Si Dengkuai, Ge Xinyu, Zeng Tao, & Li Xiaolin. (2019). Housing Price Volatility, Financial Stability, and Optimal Macroprudential Policy. *Financial Research*, 473(11), 38-56.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

AI Intervention in the Irrational Behavior of Individual Investors in the Securities Industry -- Based on the Construction of Financial Culture

Chengjun Zou*

College of Information Science and Technology, Jinan University, Guangzhou, China

**Corresponding author: Chengjun Zou.*

Abstract

In view of the practical problems of “limited effect of independent implementation and fuzzy coordination mechanism” in the current AI intervention and financial culture construction in guiding investors' rational behavior, this study takes the panel data of 31 provinces in mainland China from 2022 to 2025 as samples, takes the risk awareness search index as the proxy variable of rational cognition, and empirically tests the interaction between the two by using a two-way fixed effect model. The results show that both single AI intervention and financial culture construction have a significant negative impact on investors' risk awareness, and the interaction between the two is significantly positive. Financial culture construction can positively adjust the effect of AI intervention, forming a “technology culture” double drive promotion mechanism. This paper fills the theoretical gap between AI technology and financial culture research and provides empirical support for the construction of collaborative governance programs and the stability of the securities market.

Keywords

AI, financial culture, securities market

1. Introduction

At present, the securities market in mainland China is dominated by individual investors, accounting for more than 90%, and there are widespread irrational behaviors dominated by emotional factors, such as chasing the rise and killing the fall, frequent trading, which not only can not guarantee the rights and interests of investors, but also aggravate market fluctuations and damage the effectiveness and stability of the overall market. At the same time, the two trends of AI technology landing and regulators' vigorous promotion of the construction of financial culture are profoundly affecting the industry pattern. Intelligent investment advisers and risk warnings are gradually popularized in investment trading platforms, providing a new way to guide rational investment. The construction of financial culture also promotes the cultivation of a rational investment ecology in terms of values. However, under the changes of the two major trends, the proportion of investment losses caused by irrational behavior is still high, reflecting a significant gap between human intervention and behavior improvement.

At present, most of the existing studies analyze the guidance effect of one party, but the interaction between the two is still unclear. Xu and Wang [1] proposed the positive impact of the construction of an individual

investor culture. Zhou [2] said that AI can be used as an assistant for intelligent risk control and credit decision-making to improve the degree of behavioral rationality. Zhang [3] pointed out that AI as a consultant can reduce the investment threshold and reduce investment risk through personalized services. Eichler and Schwab [4] stressed that it is difficult to achieve long-term improvement of investors' rational decision by relying solely on AI or promoting financial education. Christina and Banerjee [5] put forward key findings, pointing out that financial literacy can regulate the impact of AI. Aristei and Gallo [6] also pointed out that there is a complex relationship between financial literacy and the use of AI counseling. This paper aims to explore the interaction between AI intervention tools and the construction of financial culture, explore the impact on the rational behavior of individual investors, reveal the mechanism of action between the two through empirical analysis, and provide new ideas for solving the bottleneck of intervention, to make up for the gap between the current intervention effect and behavior improvement.

2. Research Methods and Design

2.1 Variable Selection

The variables and their definitions selected in this paper are shown in Table 1.

Table 1: Variable description

Variable	symbol	definition
Risk awareness keyword search index	Searchit	Baidu Index risk control, position explosion, stop loss weighted by 5:3:2 and provincial population weighted
AI intervention degree	Alit	Select the Provincial Digital financial service index of Tsinghua University in mainland China
Degree of financial culture construction	FCit_scaled	The number of documents issued by the investor education column of the securities regulatory bureaus of various provinces in mainland China, weighted by population and standardized
Interaction term	Alit_FCit_interaction	The result is obtained by multiplying the degree of AI intervention by the level of financial culture development.
Control	PGDP	Annual per capita GDP of each province
	PExp	Annual per capita general public budget expenditure by province
	p	Population of each province
	internet	Internet penetration rate
	urbanization	Urbanization rate

2.2 Data Sources

This paper selects the search index provided by Baidu Index from 2022 to 2025 for the three terms of “risk control”, “position explosion” and “stop loss” in mainland China each year to synthesize search volume. Based on the Provincial Digital Financial Service Index released by Tsinghua University at the end of 2025, this paper calculates the number of documents issued by provinces in mainland China in the “investor education” column of the official website of the securities regulatory bureau from 2022 to 2025. At the same time, the provincial population is weighted after the 5:3:2 weighting of the three-term search indexes in each province, and the number of column articles in each province is weighted by the population and standardized by Z-score to eliminate the extreme value impact caused by the dimensional gap. The descriptive statistical results between variables are shown in Table 2.

Table 2: Descriptive statistical analysis

Variable	Observed value	Mean value	Standard deviation	Min	Max	Median
Searchit	124	1.074	0.910	0.000	5.266	0.840
FCit_scaled	124	0.000	1.004	-0.418	3.192	-0.304
Alit	124	95.430	10.170	78.330	121.240	96.805
AI FC interaction	124	0.000	43.864	-43.778	154.926	-0.698

2.3 Model design

This paper constructs the following regression model to comprehensively analyze the impact of AI intervention and financial culture construction on investors' rational cognition:

$$\text{Searchit} = \beta_0 + \beta_1 \text{AIit} + \beta_2 \text{FCit\_scaled} + \beta_3 (\text{AI\_FC\_interaction}) + \gamma_{\text{Controlsit}} + \epsilon_{it} \quad (1)$$

In this paper, β_1 is used to reflect the impact of AI intervention on risk perception under the remaining conditions, and β_3 is used to reflect the comprehensive impact of AI intervention and financial culture construction on investor cognition. At the same time, considering that the construction of financial culture depends on the specific environment, we will not give too much explanation to the separate β_2 , but turn to the way of measuring the degree of financial culture regulation by the level of $\beta_1 + \beta_3 \times \text{AI}$.

Compared with other existing studies, this study uses multiple linear regression to add the cross term of AI and FC to the regression equation, and uses the cross term coefficient to measure the impact on investors driven by the two together, which can make the mutual promotion mechanism of the two clearer.

3. Research Findings and Analysis

3.1 Correlation analysis

Table 3: Correlation analysis

Variable	Searchit	FCit_scaled	AIit	AI_FC_interaction
Searchit	1.000			
FCit_scaled	-0.25***	1.000		
AIit	-0.60***	0.08	1.000	
AI_FC_interaction	0.15*	0.85***	0.40***	1.000

The correlation test results of variables between samples are shown in Table 3. It can be seen that the correlation coefficient between financial culture construction level (fcit_scaled) and investors' risk awareness (searchit) is significantly negative at the level of 1%. At the same time, there is also a highly significant negative correlation between AI intervention intensity (AIIT) and investors' risk awareness (searchit).

The absolute values of the coefficients between the core explanatory variables (fcit_scaled, AIIT) and the explained variables (searchit) are less than 0.8, which preliminarily indicates that the risk of serious multicollinearity problems in the model is low. It should be noted that there is a high correlation coefficient between ai_fc_interaction and one of its constituent variables, standardized fcit_scaled, which is determined by the multiplicative construction of the interaction term and is within the theoretical expectation.

3.2 Multicollinearity

This study further calculated the variance inflation factor (VIF). The Vif values of all core explanatory variables were lower than 5 (see Table 4), which was far lower than the critical standard of 10, indicating that the core explanatory variables were relatively independent, while the Vif values of control variables were mainly between 3 and 10. This phenomenon stems from the close internal logic between economic development, digital development and the urbanization process, and the control variables did not have a substantial impact on the estimation results of the core explanatory variables, which confirmed that multicollinearity would not interfere with the reliability of the empirical results in this paper.

Table 4: Multicollinearity test

Variable	VIF	1/VIF
FCit_scaled	1.18	0.847
AIit	1.23	0.813
AI_FC_interaction	1.31	0.763
PGDP	8.76	0.114
PExp	7.82	0.128
p	3.45	0.290
urbanization	9.21	0.109
internet	8.95	0.112

3.3 Hausman Test

This study employed the Hausman test to evaluate the selection results of the research model. According to the test results, a p-value of <0.1 was obtained, leading to the rejection of the null hypothesis. Therefore, a fixed effects model was adopted.

3.4 Baseline Regression Results

Table 5 reflects the degree of impact of AI intervention and financial culture construction on investors' rational cognition, and the common impact of the two. Column 1 is the common case, column 2 is the addition of control variables, and column 3 is the addition of control variables and fixed effects. The results show that the interaction coefficient is significantly positive at the level of 1% regardless of whether the control variable and the fixed effect of the province year are added, indicating that the synergy of AI intervention and financial culture construction can significantly improve the rational cognition of investors weighted by population. At the same time, the AIIT coefficient and the FCIT coefficient show that both of them have significant negative effects on investors.

Table 5: Regression analysis

variable	1 General circumstances	2 Add control variables	3 Add fixed effects (control + two-way fixed effects)
AIit	-0.0173***(-24.578)	-0.0165***(-22.892)	-0.0152***(-20.123)
FCit scaled	-0.2880**(-3.249)	-0.2750**(-3.012)	-0.2600**(-2.876)
AI FC interaction	0.0173*** (24.578)	0.0165*** (22.892)	0.0152*** (20.123)
PGDP	-	0.3780***(-8.932)	0.3210*** (6.987)
PExp	-	0.00019***(-7.456)	0.00017***(-5.892)
P (Permanent resident population in ten thousand)		-0.00002*(-1.892)	-0.00001(-1.234)
urbanization		0.0021**(-2.345)	0.0015*(-1.923)
internet		0.0018**(-2.123)	0.0016**(-2.012)
cons	2.1799***(-24.578)	2.0120***(-19.876)	2.0150***(-21.345)
Ind	No	No	Yes
Year	No	No	Yes
N	124	124	124
R ²	0.8917	0.9245	0.9482
Adj.R ²	0.8892	0.9201	0.9445

4. Discussion

From other studies, Zou [7] affirmed the advantages provided by AI in the decision-making stage, but also pointed out that investors generally blindly believed in and followed AI's big data-based recommendations, amplified the "herding effect" to a certain extent, and believed that institutions should focus on AI content. Wu and Li [8] pointed out in their research that the intelligent algorithms of most institutions converge, and because the decision-making process is invisible, once AI makes a wrong judgment, it will cause a chain reaction in the market, affecting the independent judgment of individual investors, advocating the diversified development of AI algorithms among institutions, and making the decision-making process as public as possible. Zhang et al. [9] pointed out the significant positive role of the financial culture publicity and education base in the survey, emphasizing the impact of offline education of entities. Hong et al. [10] pointed out that a pure AI intervention would weaken risk awareness. Praveen et al. [11] emphasized that AI tools have more obvious benefits for users with low financial literacy. Generally speaking, scholars have a certain understanding of the negative impact of the current intervention form and education level, but the discussion on AI intervention and financial culture construction is relatively scattered.

Based on the comprehensive exploration of this study, the following suggestions are put forward: on the one hand, we should promote the standardized iteration of AI investment consulting technology, focus on strengthening the transparency and interpretability of the algorithm, and give a clear risk prompt for the recommendations and decisions of AI generation; On the other hand, we should vigorously carry out offline and hierarchical financial culture education and publicity activities, and build an enabling mechanism for the integration of AI and financial culture, to break the bottleneck of single intervention.

5. Conclusion

Based on the provincial panel data of mainland China from 2022 to 2025, this paper empirically tests the influence mechanism of AI intervention and financial culture construction on investors' rational cognition by using multiple linear regression under the two-way fixed effect model. The results show that the implementation of AI intervention or financial culture construction alone will have a significant negative impact on investors' risk awareness, and the interaction effect of the two is significantly positive. Financial culture construction can weaken the negative effect of AI intervention and achieve significant positive regulation. The improvement of AI intervention level can also hedge the negative effect of pure financial culture propaganda, and finally form a rational promotion mechanism driven by "technology culture". This paper verifies the core value of collaborative policy implementation, breaks through the limitation of the effect bottleneck of single technology intervention or cultural propaganda, and provides a new empirical basis and improvement logic for rational cultivation of investors.

There are still some limitations in this paper, which can be optimized and developed in the following aspects in the future. In terms of the explained variables, the search index of Baidu Index can not fully reflect the risk awareness of individual investors. With the diversification of social platforms, some investors are not excluded from querying the relevant terms in the rest. At the same time, the effect of the fixed effect model in dealing with the missing variables that change over time and are related to the core explanatory variables is not ideal, and more external factors may not be taken into account.

References

- [1] Xu, L., & Wang, Y. 2026. Analysis of Individual Investor Culture Construction in the Securities Industry under the Background of Fintech. *Market Outlook*, (1), 58-60.
- [2] Zhou, H. T. 2025. Application and Governance of Artificial Intelligence in the Financial Industry. *Financial Perspectives*, (9), 68-74.
- [3] Zhang, Y. W. 2025. What Can Robo-Advisors Bring to Personal Finance. *Cloud*, (21), 109-111.
- [4] Eichler, K. S., & Schwab, E. 2024. Evaluating robo-advisors through behavioral finance: a critical review of technology potential, rationality, and investor expectations. *Frontiers in Behavioral Economics*, 3, 1489159.
- [5] Christina, N., & Banerjee, A. 2025. Cognitive rescue or digital overreach? Artificial intelligence and the disrupted rationality of retail investors in the era of financial digital transformation. *The Business & Management Review*, 16(3), 88-98.
- [6] Aristei, D., & Gallo, M. 2025. Financial literacy, robo-advising, and the demand for human financial advice: Evidence from Italy. *Journal of Behavioral and Experimental Finance*, 101125.
- [7] Zou, Y. M. 2025. Research on the Application of Artificial Intelligence in Financial Investment Decision-making. *Sales & Management*, (32), 21-23.
- [8] Wu, H. X., & Li, J. X. 2025. Application of Artificial Intelligence in Financial Investment Decision-making. *Science Technology Economy Market*, (4), 104-106.
- [9] Zhang, T. Z., Li, B., Tan, Y. Z., et al. 2025. Chongqing Three Gorges Bank: Let Financial Services and Financial Education "Fly Side by Side" — Research Report on the Construction of Financial Culture Education and Promotion Base of Chongqing Three Gorges Bank. *Contemporary Financiers*, (10), 22-25.
- [10] Hong, C. Y., Lu, X., & Pan, J. 2025. FinTech platforms and mutual fund distribution. *Management Science*, 71(1), 488-517.
- [11] Praveen, R. V. S., Vemuri, H. K., Peri, S. S. S. R. G., Sista, S., Saxena, V., & Saxena, P. 2025. Enhancing Financial Literacy and Personal Investment Decisions Through AI and Machine Learning. *Journal of Marketing & Social Research*, 2, 268-280.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

The Dampening Effect of Green Finance Reform Policies on Corporate Green Innovation Bubbles: An Empirical Study of Chinese A-Share Listed Companies

Zixin Xu*

Sino-Russian Institute, Jiangsu Normal University, Xuzhou 221100, China

**Corresponding author: Zixin Xu.*

Abstract

While green finance reform policies channel capital toward green industries, can they effectively curb the green innovation bubbles caused by corporate “greenwashing”? This study examines Chinese A-share listed companies from 2010 to 2024, treating the establishment of green finance reform and innovation pilot zones as a quasi-natural experiment. Using a multi-period difference-in-differences method, the study systematically investigates the impact of green finance policies on corporate green innovation bubbles and their transmission mechanisms. The findings reveal: First, green finance reform policies significantly curb corporate green innovation bubbles. Second, a decomposition of the transmission mechanisms indicates that strategic green innovation is the core channel through which policies exert their “corrective” function; due to the short-term crowding-out effect of policies on substantive innovation, the contribution of the substantive green innovation channel is weak and runs in the opposite direction. Third, the policy effects exhibit heterogeneity: non-state-owned enterprises are more sensitive to policy responses; the moderating role of executives’ financial backgrounds is limited, with differences between groups failing to reach statistical significance. This paper reveals the underlying mechanism by which green finance policies improve innovation quality by curbing strategic innovation, providing empirical evidence for mitigating “greenwashing” risks and optimizing the institutional design of green finance.

Keywords

green finance reform, green innovation bubble, strategic innovation, substantive innovation

1. Introduction

Against the backdrop of the global response to climate change and China’s implementation of the “dual carbon” strategy, driving the green and low-carbon transformation of economic and social development is a key component of achieving high-quality growth. Encouraging enterprises to engage in green innovation has become a core policy option for countries seeking to foster new productive forces and address environmental challenges. To accelerate the development of a green financial system and gain a deeper understanding of its growth trajectory, the Chinese government established green finance reform and innovation pilot zones in three batches in 2017, 2019, and 2022. A large body of academic research indicates that such environmental regulations and market incentive policies can effectively stimulate enterprises’ green innovation activities,

significantly increasing their green patent output and R&D investment [1, 2]. However, driven by policy incentives, a concerning phenomenon is emerging: some enterprises, in order to align with policy directions and secure green finance support, tend to engage in low-quality, symbolic green innovation—that is, they accumulate “green” labels by filing a large number of utility model patents with low technical content, rather than pursuing substantive technological breakthroughs. This behavior has spawned a so-called “green innovation bubble,” manifested by a surge in the quantity of green patents while quality stagnates, and an abnormally high proportion of utility model patents. Once formed, this bubble not only distorts resource allocation—misdirecting limited financial resources toward “greenwashing” enterprises—but also squeezes the survival space of companies genuinely engaged in green R&D, creating a vicious cycle where “bad money drives out good.” Ultimately, this undermines corporate core competitiveness, weakens regional innovation ecosystems, and may even trigger a crisis of confidence in green industries within capital markets. Therefore, exploring whether green finance policies can effectively curb this bubble and guide enterprises toward “true green” practices holds urgent practical significance.

Research and development of green and low-carbon technologies is generally characterized by long cycles, high investment, and significant uncertainty. To some extent, this encourages companies to meet external requirements through token R&D efforts and low-quality patent applications, thereby creating a “green innovation bubble” [3]. Such “innovation bubbles” not only distort resource allocation and dilute the effectiveness of policies, but may also cause long-term damage to enterprises’ core competitiveness and the healthy development of regional innovation ecosystems. Can green finance reform policies effectively curb this bubble? What is the mechanism behind their effectiveness? Do policy outcomes vary across enterprises with different characteristics? These questions urgently require empirical testing. However, most existing studies focus on the “quantitative” effects of green innovation, while relatively little attention has been paid to examining the “quality” and “motivation” of innovation. Scholars have primarily verified the policies’ role in promoting corporate green innovation efficiency and their contribution to fostering new-quality productive forces through a focus on technological innovation [1, 4], but this overlooks the role of policy in managing financial risks, particularly bubbles at the corporate level. In the realm of innovation strategy, While Li and Zheng [5] pioneering research has provided an important perspective for understanding corporate compliance behavior, few studies have directly applied the “strategic innovation” analytical framework to the evaluation of green finance policies or examined its relationship with the “bubble” phenomenon in green innovation.

In light of this, this paper treats the establishment of Green Finance Reform and Innovation Pilot Zones as a quasi-natural experiment. Using data on A-share listed companies from 2010 to 2024 and employing a multi-period difference-in-differences method, we investigate whether this policy can curb green innovation bubbles among firms. Here, the green innovation bubble is defined as the proportion of green utility model patent applications out of the total number of green patent applications; a higher value of this indicator suggests that firms are more inclined to engage in low-quality, strategic green innovation. The study finds the following: First, the Green Finance Reform and Innovation Pilot Zone policy significantly curbs corporate “green innovation bubbles,” reducing the proportion of bubble levels among firms in the treatment group by approximately 3.9%. This conclusion remains valid after a series of robustness tests. Second, this paper further examines the policy’s mechanism of action and finds that it primarily reduces bubbles by curbing firms’ strategic green innovation; this pathway accounts for 69.5% of the total effect; Third, heterogeneity analysis reveals that the policy’s dampening effect is more pronounced in non-state-owned enterprises and firms with executives possessing a financial background. This is because, in these firms, the policy exerts a stronger inhibitory effect on strategic innovation. Additionally, the study finds that the policy generates a short-term crowding-out effect on substantive green innovation, which to some extent weakens the policy’s net effect and offers insights for future policy optimization.

The marginal contributions of this paper are primarily reflected in the following four aspects: First, in terms of research perspective, unlike previous studies that focused on the “quantity” of green innovation while paying less attention to its “quality” and “motivation,” this paper introduces the concept of “green innovation bubbles” into the field of green innovation policy impact assessment. Building on this, the paper further distinguishes between the differential impacts of strategic green innovation and substantive green innovation on green financial bubbles, finding that strategic innovation significantly exacerbates bubbles, whereas substantive innovation suppresses them. This finding not only enriches and deepens our understanding of the micro-level effects of green innovation incentive policies from a new perspective of innovation quality and intrinsic value,

but also reveals the potential harm of “greenwashing” from the perspective of economic consequences, highlighting the necessity of policy intervention and providing supplementary evidence regarding the effectiveness of green finance policies from the perspective of capital markets. Second, regarding the mechanism of action, this paper breaks away from the traditional framework of financing constraints and the intensity of environmental regulations to systematically reveal the transmission channels through which policies influence green innovation bubbles. It confirms for the first time that “strategic green innovation” is the core micro-mechanism of policy transmission, accounting for 69.5% of the total effect; Third, regarding policy implications, this study reveals the “inclusive” nature of green finance reform policies in risk governance. That is, their effectiveness does not depend on local government fiscal investment or the intensity of environmental enforcement, but rather can penetrate regional institutional barriers through channels such as credit allocation and market constraints to directly influence firm-level behavior. At the same time, the short-term crowding-out effect of the policy on substantive innovation suggests that adjustment costs must be considered when guiding corporate transformation, providing empirical support for optimizing policy design and establishing a unified national green finance system.

The remainder of this paper is organized as follows: Section 2 introduces the policy background; Section 3 presents a literature review and theoretical analysis; Section 4 outlines the empirical research design; Section 5 reports the empirical results and analysis; Section 6 provides further analysis; and Section 7 summarizes the conclusions and offers corresponding policy recommendations.

2. Policy Background

As noted in the introduction, while green finance directs capital toward green industries, it also faces potential risks such as “greenwashing” and “green innovation bubbles.” How to design effective incentive and constraint mechanisms that both leverage finance’s role in driving the green transition and guard against market distortions caused by corporate strategic innovation has become a critical challenge for policymakers. To thoroughly explore a green finance development path suited to China’s national conditions, the country has adopted a reform strategy of “piloting first, then gradual rollout,” using regional trials to provide empirical evidence for institutional design.

To thoroughly explore a green finance development path suited to China’s national conditions and establish effective incentive and constraint mechanisms, China has adopted a reform strategy of “piloting first, then gradually expanding.” The issuance of the “Guiding Opinions on Building a Green Finance System” marked the establishment of a policy guidance framework in the field of green finance. In 2017, China established green finance reform and innovation pilot zones in regions such as Zhejiang, Jiangxi, Guizhou, Guangdong, and Xinjiang. In 2019 and 2022, the state respectively included the Lanzhou New Area in Gansu and Chongqing Municipality in the pilot scope, forming a policy implementation pathway of “piloting first, gradual rollout.” The selection of these pilot zones was not arbitrary but followed clear principles aimed at generating replicable and scalable experiences: First, the principle of geographical representativeness, ensuring that the pilots cover regions with varying levels of economic development, including the eastern, central, western, and northeastern regions; Second, the principle of resource endowment diversity, which fully considers the unique characteristics of each region in terms of ecological resources and industrial foundations; third, the principle of resource endowment diversity, which encourages localities to explore differentiated green transition models based on their own strengths. For example, Zhejiang Province focuses on utilizing financial instruments to promote the green upgrading of traditional manufacturing and the innovation of financial products; Guangdong Province, leveraging its developed financial markets, focuses on exploring the integration of carbon finance with the new energy industry; Guizhou Province is committed to promoting the convergence of big data, green industries, and green finance; and Gansu Province (Lanzhou New Area) emphasizes the application of green finance in ecological conservation within the Yellow River Basin and in inclusive finance. The diverse practices in these pilot regions provide a rich institutional context for this paper to examine the micro-level effects of green finance policies at the enterprise level.

It is worth noting that while advancing green finance, pilot regions have also begun to address the risks posed by strategic corporate innovation. Some regions have incorporated measures such as environmental disclosure requirements, green credit review mechanisms, and differentiated interest rate pricing into their policy designs, effectively establishing institutional constraints against “greenwashing.” As a form of

organizational environmental misconduct, greenwashing essentially involves constructing an eco-friendly image through systematic false advertising while evading substantive environmental responsibility and investment. Externally, environmental subsidies, regulatory pressure, regulatory loopholes, and consumer demand for green products collectively create incentives for greenwashing, while information asymmetry provides the space for its implementation. Internally, factors such as low visibility, organizational inertia, the pursuit of environmental reputation, and external financing pressures also drive companies to engage in greenwashing; this tendency is particularly pronounced among highly indebted firms. Research by He et al. [6] indicates that strengthening external oversight can effectively curb corporate greenwashing; for example, central environmental inspections have significantly reduced the extent of greenwashing among polluting enterprises. Therefore, by embedding constraint mechanisms such as information disclosure and credit review within green finance systems, pilot regions are effectively curtailing the scope for greenwashing from an external regulatory perspective. These initiatives provide a rich institutional context for this paper to examine the governance effects of green finance policies on innovation bubbles at the corporate level. By evaluating policies in pilot zones, we can not only verify the effectiveness of the policies themselves but also reveal the unique role of market-based tools in preventing strategic innovation, thereby providing empirical support for the development of a unified national green finance system.

3. Literature Review and Theoretical Hypotheses

3.1 Literature Review

Given that this study examines the factors influencing whether green finance reform and innovation pilot zones can curb green innovation bubbles in enterprises—and considering that green innovation bubbles essentially represent a concentrated manifestation of strategic innovation behavior by enterprises in the green sector—this paper will conduct a literature review focusing on two aspects: strategic innovation and green innovation bubbles.

3.1.1 Strategic Innovation

Identifying the motivations behind corporate innovation is a central issue in the study of the economics of innovation. Li and Zheng's [5] pioneering research systematically distinguishes between substantive innovation and strategic innovation for the first time: substantive innovation refers to “high-quality” innovative activities aimed at driving technological progress within firms and gaining a competitive advantage; strategic innovation, on the other hand, refers to innovation strategies designed to pursue other benefits by prioritizing the “quantity” and “speed” of innovation in order to comply with regulatory requirements and government incentives. This distinction overcomes the limitations of traditional innovation research, which focuses solely on total patent volume, and provides an important analytical tool for understanding the “innovation distortions” that may arise in firms under policy incentives.

Subsequent research has further expanded this framework. When firms face environmental regulations, strategic green innovation may also occur, as Shen et al. [7] have observed, local governments tend to prioritize the quantity of patents over their quality when supporting innovative enterprises. In the area of tax incentives, Yang and Rui [8] found that enterprises certified as high-tech enterprises can obtain more tax breaks and government subsidies. The tax reductions, government subsidies, and talent aggregation effects resulting from industrial policies only promote increased innovation among genuine high-tech enterprises, whereas the incentive effect of industrial policies on innovation is significantly weakened for “pseudo-high-tech enterprises.” In the field of green credit, Liu [9] drawing on China's economic transformation practices, conducted an empirical examination of the impact of green finance policies on firms' strategic green innovation behavior and found that firms exhibit a tendency to engage in strategic green innovation by leveraging green finance policy support.

3.1.2 Green Innovation Bubble

The concept of the “innovation bubble” stems from observations of a divergence between the quantity and quality of innovation. In the field of green innovation, the “green innovation bubble” refers to a phenomenon in which green innovation activities experience rapid quantitative growth driven by multiple factors—including policy incentives, capital flows, and market signals—yet their quality and efficiency fail to improve

at the same pace, resulting in resource misallocation and inflated valuations [10]. A key indicator of this trend is the unusually high proportion of green utility model patents, as these patents do not undergo substantive examination and are therefore more easily used by companies as strategic tools.

This paper defines a “green innovation bubble” as a structural imbalance in which, while the number of green patents held by enterprises increases, the proportion of low-quality patents (utility model patents) rises abnormally. The core metric is the proportion of green utility model patents out of the total number of green patents. The higher this indicator, the more “inflated” the enterprise’s green innovation appears, and the more severe the bubble.

The formation mechanism of the green innovation bubble can be explained from multiple perspectives. First, the combination of policy incentives and regulatory gaps. Fiscal support is insufficient, and effective regulatory measures are lacking. Some local governments have introduced incentives such as fiscal subsidies and tax breaks to meet environmental assessment targets; however, due to the absence of effective evaluation and oversight, enterprises exploit a large number of low-quality green patents to obtain policy resources, resulting in the widespread phenomenon of “nominal innovation.” Second, the transmission of speculative behavior in capital markets. Irrational speculative activities in the stock market have exacerbated the risk of market bubbles. Against a backdrop of widespread social responsibility awareness, massive capital inflows into the environmental protection sector have caused the valuation of related companies to deviate from their actual innovation capabilities and profit levels. Third, information asymmetry regarding green technologies. Due to the unique nature of green technologies and information asymmetry, investors and regulators find it difficult to accurately assess their true value, leading to the phenomenon of “greenwashing.” In addition to micro-level factors, the macro-level institutional environment is also a key factor influencing the formation and development of green innovation bubbles. Beyond micro-level factors, the macro-level institutional environment similarly impacts the formation and development of green innovation bubbles. The national unified market serves as a crucial institutional platform for promoting the flow of factors of production, fostering fair competition, and eliminating trade barriers between regions. It helps improve the efficiency of green technology allocation across China’s regions while reducing information barriers and resource misallocation. Advancing this process can not only facilitate the flow of innovation resources and enhance green innovation but also effectively curb innovation bubbles caused by policy arbitrage and market closure. However, existing research lacks systematic analysis of how China’s unified national market influences the formation of green innovation bubbles, and there is an urgent need for in-depth exploration from an integration perspective.

3.1.3 Research Gaps

In summary, existing research leaves room for further exploration in the following areas: First, few studies have directly linked green finance policies to “green innovation bubbles” (rather than simply the quantity of innovation or strategic innovation itself) to examine the governance effects of policies on the quality structure of innovation. Second, the underlying mechanisms through which policies influence firms’ strategic innovation remain unclear and have not been systematically tested empirically. Third, the differentiated impacts of strategic innovation and substantive innovation on the same economic outcomes (such as firm value and innovation bubbles) have not yet been fully elucidated.

3.2 Theoretical Analysis and Research Hypotheses

Within the important institutional framework of China’s green finance policies, enterprises constitute a vital component of the country’s economic development and serve as a key driving force in promoting economic growth and achieving a green transition. Therefore, examining the significant role of China’s green finance pilot policies in national economic development from the perspective of enterprises holds important theoretical value for the formulation and implementation of relevant policies.

3.2.1 The General Effect Hypothesis

In markets characterized by information asymmetry, companies—as parties possessing private information about their own green technological capabilities—need to convey their “green credentials” to principals such as governments, investors, and financial institutions by sending observable signals (such as green patents). Different types of patents serve as signals with significant variations in cost and credibility: substantive green

innovation (invention patents) involves long R&D cycles, substantial investment, and high technical barriers, making it a high-cost signal with high credibility; strategic green innovation (utility model patents) entails low R&D costs and rapid authorization, making it a low-cost signal with low credibility. In the absence of an effective signal screening mechanism, the market may fall into a “convergence equilibrium”—where all firms send low-cost signals (strategic innovation), leading to “Gresham’s law” and the formation of a green innovation bubble. The root cause of this agency problem lies in the fact that management, acting as corporate agents, may be motivated by self-interest to pursue rapid growth in patent volume to quickly capture policy dividends and boost short-term performance, while neglecting technological substance and environmental benefits.

The introduction of policies for Green Finance Reform and Innovation Pilot Zones has fundamentally altered the rules of this signaling game. As mentioned earlier, green finance policies, by directing financial resources toward enterprises with genuine green technological capabilities, effectively raise the implicit costs of firms sending low-cost signals (strategic green innovation) and enhance the returns on high-cost signals (substantive green innovation). On the one hand, green finance policies typically come with strict criteria for green project certification and environmental disclosure requirements. This makes it difficult for “pseudo-green” enterprises—those relying solely on strategic patent packaging—to pass reviews and obtain financing support such as green loans and green bonds, significantly increasing the cost of their low-cost “disguise.” On the other hand, for firms with substantive green innovations, green finance policies provide positive incentives such as lower financing rates, more accessible financing channels, and government subsidies, thereby amplifying the returns on high-cost signaling. This reshaping of the cost-benefit structure will disrupt the existing “equilibrium of conflation,” prompting firms to shift from strategic innovation focused on short-term patent volume to substantive innovation emphasizing technological quality, thereby curbing the formation of a green innovation bubble. Based on this, this paper proposes the main hypothesis:

H₁: Policies in green finance reform and innovation pilot zones help curb green innovation bubbles among enterprises.

3.2.2 Mechanistic hypothesis

Furthermore, this paper analyzes the mechanisms through which green finance reform pilot zones exert their dampening effect on corporate green innovation bubbles. The policy reduces the marginal cost of substantive innovation through positive incentives, while simultaneously increasing the marginal cost of strategic innovation through negative constraints; the combined effect of these two mechanisms leads firms to reduce strategic innovation activities. On the other hand, strategic innovation is the direct source of green innovation bubbles—when firms engage in low-quality, short-sighted strategic innovation, it directly leads to an increase in the proportion of green utility model patents, resulting in an “inflated” innovation structure. Based on this, the following hypotheses are proposed:

H_{2a}: Policies in green finance reform and innovation pilot zones help curb strategic green innovation by enterprises.

H_{2b}: Strategic green innovation may exacerbate the green innovation bubble among companies.

4. Empirical Strategies and Data

4.1 Sample Data

This study uses Chinese A-share listed companies from 2010 to 2024 as the research sample. Green patent data was obtained from the China Research Data Service Platform (CNRDS), corporate financial and governance data from the Guotai-An Database (CSMAR), and the list of green finance reform pilot zones was manually compiled based on relevant policy documents issued by the People’s Bank of China and other authorities. Following existing research, the initial sample was screened as follows: (1) companies classified as ST, *ST, or SST, as well as those delisted during the study period, were excluded; (2) financial sector listed companies were excluded; (3) companies listed for less than three years were excluded. To eliminate the impact of outliers, all continuous variables were trimmed at the 1st and 99th percentiles. This resulted in a final dataset of 34,278 company-year observations, covering 3,919 listed companies distributed across 271 cities.

4.2 Model Setup

To examine the impact of policies in the green finance reform pilot zone on green innovation bubbles among enterprises, this paper constructs a multi-period difference-in-differences model as follows (main model):

$$GBR_{i,t} = \beta_0 + \alpha_1 DID_{i,t} + \gamma Controls_{i,t} + \delta_i + \lambda_t + \theta_c + \varepsilon_{i,t} \quad (1)$$

Here, i, c, t represent firm, city, and year, respectively; GBR is the dependent variable; DID is the key independent variable; $Controls_{i,t}$ are control variables; δ_i is the firm-level fixed effect, λ_t is the year-level fixed effect, θ_c is the city-level fixed effect; $\varepsilon_{i,t}$ is the random error term. Standard errors are cluster-adjusted at the city level to address within-group correlation issues. If α_1 is significantly negative, it indicates that the policy significantly suppressed the green innovation bubble among firms, and hypothesis H_1 is supported.

4.3 Variable Definitions

4.3.1 Dependent Variable

Green Innovation Bubble (GBR). Drawing on existing research, this paper defines the green innovation bubble as the proportion of green utility model patent applications out of the total number of green patent applications. A higher value for this indicator suggests that a larger proportion of patents in corporate green innovation are of lower quality, and that the bubble is more severe. Green patents are identified based on the green list of the International Patent Classification published by the World Intellectual Property Organization (WIPO).

4.3.2 Key Explanatory Variable

Policy dummy variable (DID). This paper employs a multi-period difference-in-differences (DID) model to evaluate the policy effects. If a firm's city is designated as a green finance reform pilot zone in year t or thereafter, the DID takes a value of 1 in year t and subsequent years; otherwise, it is 0. China established green finance reform and innovation pilot zones in three batches in 2017, 2019, and 2022. This paper generates the DID variable based on the pilot year corresponding to the firm's city.

4.3.3 Mechanism Variable

Strategic Green Innovation (TGU): Drawing on the research by Li Wenjing and Zheng Manni (2016), this paper uses the number of green utility model patent applications as a proxy for strategic green innovation. This variable refers to the number of green utility model patents for which a company filed applications independently in a given year.

Substantial Green Innovation (TGI): Use the number of green invention patent applications as a proxy for substantive green innovation. This variable represents the number of green invention patents for which the company filed applications independently during the year.

4.3.4 Control Variable

Drawing on existing research, this study selects the following characteristic variables that may influence corporate green innovation as control variables: *Firm Size*, *Debt To Asset*, *ROA*, *Current Ratio*, *Top1 holder*, *Tobins Q*, *Inst holding*, *L rd ratio*. In addition, this paper controls for firm-level, city-level, and year-level fixed effects. The specific variable definitions are as follows: Table 1.

Table 1: Variable Declaration

Variable type	Variable name	Variable Definition
Dependent variable	<i>GBR</i>	Percentage of green utility model patents
Key explanatory variable	<i>DID</i>	If the city where the company is located is a pilot city in the current year and subsequent years, the value is 1; otherwise, it is 0.
Mechanism variable	<i>TGU</i>	Total number of green utility model patents
	<i>TGI</i>	Total number of green invention patents
Control variable	<i>Firm Size</i>	The natural logarithm of total assets at year-end
	<i>Debt To Asset</i>	Total Liabilities at Year-End / Total Assets at Year-End
	<i>ROA</i>	Net income / Total assets

Variable type	Variable name	Variable Definition
	<i>Current_Ratio</i>	Current assets / Current liabilities
	<i>Top1_holder</i>	Number of shares held by the largest shareholder / Total number of shares
	<i>Tobins_Q</i>	Market value / Replacement cost
	<i>Inst_holding</i>	Number of shares held by institutional investors / Total number of shares
	<i>L_rd_ratio</i>	R&D expenses / revenue for the prior period
Heterogeneous variable	<i>SOE</i>	Return 1 if it is a state-owned enterprise; otherwise, return 0
	<i>FinBack</i>	If members of the board of directors, supervisors, and senior management have a financial background, assign a value of 1; otherwise, assign a value of 0

4.3.5 Descriptive Statistics

Table 2 presents the descriptive statistics for the main variables.

Table 2: Descriptive Statistics

Variable name	Mean	Standard deviation	Minimum	Maximum	N
<i>DID</i>	0.031	0.174	0	1	34278
<i>GBR</i>	0.118	0.282	0	1	34278
<i>TGU</i>	0.547	1.755	0	17	34278
<i>TGI</i>	0.757	2.475	0	30	34278
<i>Firm Size</i>	22.150	1.183	19.867	26.314	34278
<i>Debt To Asset</i>	0.410	0.196	0.057	0.914	34278
<i>ROA</i>	0.410	0.055	-0.242	0.184	34278
<i>Current_Ratio</i>	2.486	2.165	0.319	15.364	34278
<i>Top1_holder</i>	0.331	0.140	0.084	0.727	34278
<i>Tobins_Q</i>	1.978	1.080	0.834	7.998	34278
<i>Inst_holding</i>	0.412	0.241	0.005	0.907	34278
<i>L_rd_ratio</i>	0.048	0.047	0.0002	0.314	34278

Note: Continuous variables have been truncated at the 1% and 99% levels.

5. Empirical Findings and Analysis

5.1 Evaluation of the Parallel Trends Hypothesis

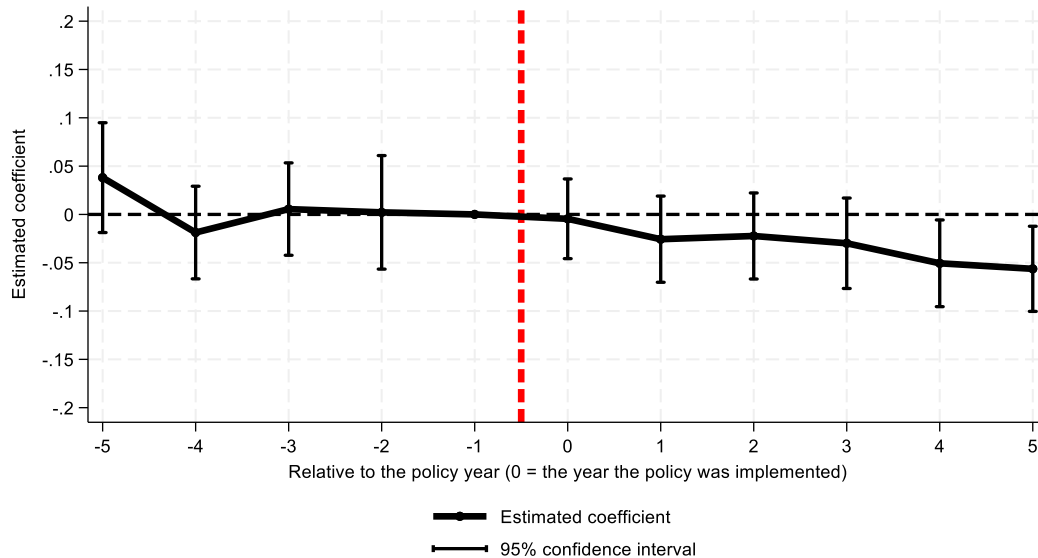
Before applying the multiple-time-point double-difference model, it is necessary to verify whether the treatment and control groups satisfy the parallel trends assumption prior to policy implementation; that is, before the implementation of policies in the Green Finance Reform and Innovation Pilot Zones, there should be no systematic differences in the degree of green innovation bubbles between enterprises within the pilot zones and those outside them. Following the approach in the existing literature, this study employs the event study method to test for parallel trends and constructs the following model:

$$GBR_{i,t} = \alpha + \sum_{k=-5}^5 \beta_k \times D_{it}^k + \gamma Controls_{it} + \delta_i + \lambda_t + \theta_c + \varepsilon_{it} \quad (2)$$

Here, D_{it}^k denotes Firm idummy variable for the k th year before and after policy implementation in the city where firm i is located, with the period immediately preceding policy implementation ($k = -1$) serving as the baseline. Specifically, using the period immediately preceding policy implementation as the baseline, we examine the differences in green innovation bubbles between firms in the treatment group and the control group across periods before and after policy implementation. As shown in Figure 1, the estimated coefficients for all periods prior to policy implementation are not significantly different from 0, indicating that the degree of green innovation bubbles among enterprises in the pilot zone before policy implementation did not differ significantly from that of enterprises outside the pilot zone, passing the parallel trends test. Furthermore, the policy effect exhibits a lag of approximately three years, becoming significantly negative starting from the fourth year after policy implementation, with the effect's intensity gradually increasing over time. This dynamic characteristic aligns with signaling theory: the policy operates by strengthening signal discrimination criteria and altering the relative costs of signaling, requiring firms time to perceive policy signals and adjust their innovation strategies. As the policy continues to be implemented, the signal discrimination mechanism is

continuously strengthened, and the cost of disguising strategic innovation continues to rise; consequently, the policy effect exhibits a cumulative strengthening pattern. The policy's corrective function is sustainable in the long term; coefficients in subsequent periods after policy implementation gradually turn negative and become significant, indicating that the policy significantly suppressed the green innovation bubble among firms in the pilot zone, which is consistent with the main conclusion of this paper. Figure 1 reports the results of the parallel trends test.

Figure 1: Evaluation of the Parallel Trends Hypothesis



Note: Period -1 serves as the base period. Insignificant pre-policy coefficients support the parallel trends assumption.

5.2 Regression to the Mean

To test the impact of policies in green finance reform pilot zones on green innovation bubbles in firms (hypothesis H_1), this study conducts estimations by progressively incorporating fixed effects and control variables. Table 3 presents the regression results. Column (1) presents a mixed OLS regression containing only the core explanatory variables and the DID; Column (2) controls for city and year fixed effects on top of this; Column (3) includes control variables but does not control for firm fixed effects; Column (4) controls for firm, city, and year fixed effects as well as control variables, and serves as the benchmark model for this study. Table 3 reports the benchmark regression results.

The results show that the coefficient of the DID in Column (1) is -0.027, which is significant at the 1% level, suggesting that the policy has a dampening effect on the green innovation bubble. In Column (2), after incorporating city and year fixed effects, the absolute value of the DID coefficient increases to -0.038 and remains highly significant, indicating that the policy effect becomes more pronounced after controlling for city- and year-specific macroeconomic factors. In Column (3), which includes control variables but does not account for firm-level fixed effects, the DID coefficient is -0.031 and is significantly negative; however, the adjusted R^2 is only 0.098, indicating limited explanatory power of the model. In Column (4), firm-level fixed effects are further controlled for. The DID coefficient is -0.039, significant at the 1% level, and the adjusted R^2 increases substantially to 0.316, indicating that firm-level heterogeneity is a key factor explaining differences in green innovation bubbles. Therefore, this paper adopts Column (4) as the baseline model.

From an economic perspective, the mean value of the green innovation bubble is 0.118. The DID coefficient of -0.039 in the baseline model implies that, following policy implementation, the proportion of the green innovation bubble among firms in the treatment group decreased significantly by approximately 3.9 percentage points compared to the control group—equivalent to 35.1% of the mean. This indicates that the policy effect possesses both statistical and economic significance. In summary, **H_1 has been proven.**

Table 3: Regression to the Mean: The Impact of Policy on the Green Innovation Bubble

Variable name	(1)	(2)	(3)	(4)
	<i>GBR</i>	<i>GBR</i>	<i>GBR</i>	<i>GBR</i>
<i>DID</i>	-0.027***	-0.038***	-0.031***	-0.039***
	(0.007)	(0.014)	(0.008)	(0.014)
<i>Firm Size</i>			0.009**	0.009*
			(0.004)	(0.005)
<i>Debt To Asset</i>			0.059***	-0.012
			(0.019)	(0.019)
<i>ROA</i>			0.241***	0.024
			(0.042)	(0.038)
<i>Current Ratio</i>			-0.002	-0.002
			(0.001)	(0.001)
<i>Top1 holder</i>			0.050***	0.044
			(0.019)	(0.031)
<i>Tobins Q</i>			-0.007***	-0.001
			(0.003)	(0.002)
<i>Inst holding</i>			-0.052***	-0.019
			(0.013)	(0.022)
City fixed effects		Yes		Yes
Year fixed effects		Yes		Yes
Firm fixed effects		Yes		Yes
N	34124	34124	34124	34016
R ²	0.000	0.019	0.008	0.316

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. The figures in parentheses represent robust standard errors aggregated to the city level. The same applies below.

5.3 Stability Test

5.3.1 Change Cluster Level

In the baseline regression, standard errors were clustered at the industry-city-industry level. To further test the sensitivity of the findings to the choice of clustering level, this study re-estimated the model with standard errors clustered at the city-industry level. This level accounts for potential within-group correlations among firms within the same industry in the same city (such as local industrial policies and regional agglomeration effects). The results in Column (1) of Table 4 show that the DID coefficient is -0.039, which is significant at the 5% level (standard error = 0.018). This is highly consistent with the results of the baseline regression, indicating that the core conclusions are not affected by the choice of clustering level.

5.3.2 Exclude Samples Collected During the Pandemic

The COVID-19 pandemic from 2020 to 2022 may have had a significant impact on corporate innovation activities. To eliminate the confounding effects of this exceptional event, this study excluded the 2020–2022 samples and reran the regression analysis. The results in Column (2) of Table 4 show that the coefficient of the DID is -0.044, which is significant at the 1% level (standard error = 0.010). This is highly consistent with the results of the baseline regression, indicating that the core conclusions are not affected by the choice of clustering hierarchy.

5.3.3 Control for Other Policy Factors

During the study period, other concurrent policies may have influenced firms' green innovation. Building on the baseline model, this paper sequentially incorporates the low-carbon city pilot policy and the innovative city pilot policy as additional control variables. The results are reported in columns (3), (4), and (5) of Table 4. After controlling for these policies, the DID coefficient remains significantly negative at the 1% level, with a stable magnitude of -0.040. Meanwhile, the coefficients of the two policy variables themselves are not significant, indicating that the core conclusion of this paper is not affected by interference from other policies and exhibits a high degree of robustness. The policy variable for the Ecological Civilization Pilot Demonstration Zone was automatically omitted by the model due to multicollinearity with fixed effects. This reflects, to some extent, that the effects of this policy have been absorbed by unobservable factors at the city

level. Nevertheless, the coefficient of the core explanatory variable, DID, remains significant and stable across all models, indicating that the conclusions of this study are not subject to interference from other policies. Table reports the results of a series of robustness tests.

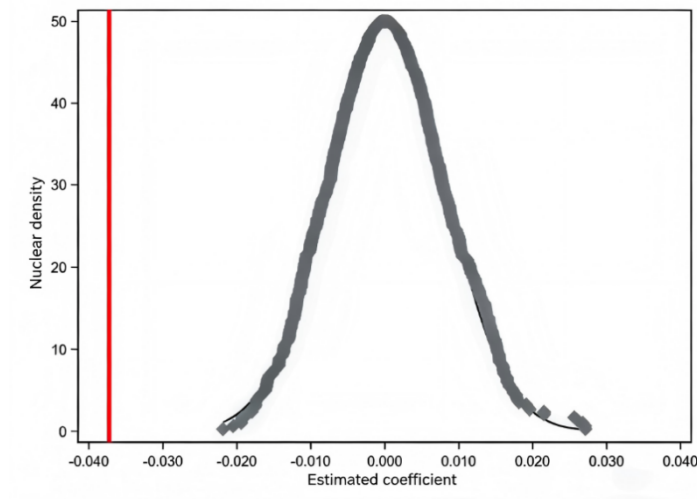
Table 4: Stability Test Results

Variable name	(1)	(2)	(3)	(4)	(5)
	City-Industry Clustering	Exclude COVID-19	Filter by Low-Carbon City Pilot Policies	Filter by Innovative Pilot Policies	Filter by All
<i>DID</i>	-0.039**	-0.044***	-0.040***	-0.040***	-0.040***
	(0.018)	(0.010)	(0.014)	(0.014)	(0.014)
<i>Lowcarbon</i>			-0.010		-0.010
			(0.009)		(0.009)
<i>Innovcity</i>				-0.003	-0.008
				(0.017)	(0.014)
Control variables	Yes	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes	Yes
N	34016	24960	34016	34016	34016
R ²	0.322	0.299	0.316	0.316	0.316

5.3.4 Placebo Test

To rule out the interference of random factors on policy effects, this study also conducts a placebo test by randomly selecting samples from the experimental group. Following the approach of Lu et al. [11], and while keeping the policy implementation period constant, this study randomly samples firms from the sample based on the number of treatment groups in the baseline regression, designating them as a “pseudo” treatment group (i.e., firms within a virtual green finance reform and innovation pilot zone). Specifically, we randomly generated 1,000 combinations of treatment and control groups and varying years, re-estimating the model after each random sampling to obtain 1,000 “pseudo” estimated coefficients. As shown in Figure 2, the 1,000 estimated coefficients obtained from the random sampling are concentrated around the origin. The red vertical line in the figure represents the true estimated coefficient (-0.039) obtained from the baseline regression in this study, which is clearly located at the left tail of the overall distribution. Therefore, the results of the baseline regression in this study are unlikely to be due to chance, which indirectly supports the robustness of these results. Figure 2 reports the results of the placebo test.

Figure 2: Placebo Test Results



6. Further Analysis

6.1 Mechanism Analysis

Drawing on Shi Lei et al.'s [12] “action pathway decomposition” framework, this paper examines how green finance reform policies influence corporate green innovation bubbles through two possible pathways: the strategic green innovation pathway and the substantive green innovation pathway. This section first tests the significance of each pathway separately, and then quantifies the contribution share of each pathway.

6.1.1 Mechanism Test: Strategic Green Innovation

Based on the theoretical framework discussed earlier, policies in green finance reform and innovation pilot zones may curb the green innovation bubble by suppressing firms' strategic green innovation and encouraging substantive green innovation (with the former being the primary mechanism). Therefore, this paper replaces the dependent variable in the main model with the mechanism variable; the specific results are shown in columns (1) and (2) of Table 5.

Table 5: Regression Results of the Mechanism Test

Variable name	(1)	(2)
	<i>TGU</i>	<i>TGI</i>
<i>DID</i>	-0.319***	-0.129*
	(0.048)	(0.069)
Control variables		
City fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Firm fixed effects	Yes	Yes
N	34066	34016
R ²	0.467	0.317

The results in Column (1) show that the DID coefficient for strategic green innovation is -0.319, which is significant at the 1% level, indicating that the policy significantly suppresses firms' strategic green innovation; **H_{2a} is thus supported.**

The results in Column (2) show that the DID coefficient for substantive green innovation is -0.129, which is significant at the 10% level, indicating that the policy has a crowding-out effect on substantive green innovation in the short term. Further regression analysis reveals that substantive innovation itself has a significant negative impact on the green innovation bubble (coefficient: -0.0046, $p < 0.01$). However, due to the crowding-out effect of the policy, the net contribution of this channel is very small (approximately 0.0006, accounting for only 1.5%) and is not statistically significant. This result is provided as supplementary information and is not included in the hypothesis testing.

6.1.2 Causal Path Analysis of Policy Mechanisms

To further examine the specific mechanisms through which digital industrial cluster policies operate, this paper draws on the average treatment effect decomposition approach proposed by Zhou and Yamamoto [13]. Within the Rubin framework, we analyze how policies in green finance reform and innovation pilot zones reduce corporate green innovation bubbles through the mechanisms of strategic green innovation and substantive green innovation.

Let the total policy effect be α_I (from the main model). First, we test the effect of the policy on strategic green innovation, yielding coefficient β_1 ; second, we test the effect of strategic green innovation on the green innovation bubble (controlling for the policy variable), yielding coefficient γ_1 . Similarly, we test the effect of the policy on substantive green innovation, yielding coefficient δ_1 , and test the effect of substantive green innovation on the green innovation bubble, yielding coefficient θ_1 . The formulas for calculating the contributions of each path and the direct effect are as follows:

Contribution of the strategic innovation path: $\beta_1 \times \gamma_1$

Contribution of the substantive innovation path: $\delta_1 \times \theta_1$

Direct effect: $\alpha_I - (\beta_1 \times \gamma_1 + \delta_1 \times \theta_1)$

The proportion of each path's contribution is calculated as follows:

Proportion of strategic innovation path = $\frac{\beta_1 \times \gamma_1}{\alpha_I} \times 100\%$

Proportion of the substantive innovation path = $\frac{\delta_1 \times \theta_1}{\alpha_I} \times 100\%$

Proportion of the direct effect = $\frac{\alpha_I - (\beta_1 \times \gamma_1 + \delta_1 \times \theta_1)}{\alpha_I} \times 100\%$

If $\beta_1 \times \gamma_1$ has the same sign as α_I and accounts for a high proportion, this indicates that strategic innovation is a key pathway through which policy exerts its influence, thereby confirming Hypothesis H_{2b}.

The calculated contribution ratios for each path are presented in Table 6.

Table 6: Path Analysis: Core Path Contributions to Strategic Green Innovation

Pathways of Green Finance Reform Pilot Policies	Coefficient Value	Percentage (%)
Overall Average Effect	-0.039***	100.00%
Green Finance Reform Pilot Policies → Strategic Innovation ↔ Green Innovation Bubble	-0.027***	69.5%
Green Finance Reform Pilot Policies → Substantive Innovation ↔ Substantive Innovation	0.0006	-1.5%
Substantive Innovation → Green Innovation Bubble	-0.012	32.0%

Note: The average total effect is derived from the baseline regression. The contribution of the strategic innovation pathway is calculated as the product of the strategic innovation coefficient (-0.319) in column (1) of Table 5 and the strategic innovation coefficient (0.085) in column (1) of Table 6. The contribution of the substantive innovation path is the product of the coefficient for substantive innovation in column (2) of Table 5 (-0.129) and the coefficient for substantive innovation in column (2) of Table 6 (-0.005). The direct effect is the total effect minus the sum of the contributions from the two paths. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively (the significance of the strategic innovation path is calculated using the Delta method, $z = -6.57$, $p < 0.001$). Among these, the contribution of the substantive innovation path is positive and opposite in direction to the total effect; the percentage shown in the table is the absolute value, and its actual effect is to weaken the total effect.

As shown in Table 6, the contribution of the strategic green innovation path is -0.027, accounting for 69.5% of the total effect, and is highly significant. This indicates that strategic green innovation exacerbates green innovation bubbles in firms, confirming Hypothesis H_{2b}. The contribution of the substantive green innovation path is 0.0006, accounting for -1.5% (in the opposite direction), and is not statistically significant. This result suggests that, in the short term, the policy exerts a certain crowding-out effect on substantive green innovation, offsetting the bubble-suppressing role it would otherwise play. However, the contribution of this pathway is negligible and does not affect the dominant effect exerted by the policy through the strategic innovation pathway. The direct effect accounts for 32.0%, indicating that the policy likely operates through other

unobserved channels (such as market expectations or reputational effects) or signals that the two pathways fail to fully capture.

In summary, green finance reform policies primarily reduce green innovation bubbles by curbing firms' strategic green innovation, while the slight negative effect of the substantive innovation channel reflects adjustment costs during the policy transition period and does not affect the core conclusion.

6.2 Heterogeneity Analysis

There may be differences in the ability of firms with different characteristics to receive and respond to policy signals, which determines the heterogeneity of policy outcomes.

(1) Dimension of Ownership Structure

State-owned enterprises and non-state-owned enterprises exhibit significant differences in terms of resource acquisition, market sensitivity, and incentive-constraint mechanisms. According to Fang [14], state-owned enterprises, relying on their inherent ties to the government, possess inherent advantages in areas such as access to financing and policy support, and are relatively less sensitive to financial market signals. When policy adjustments alter the rules of the signaling game, state-owned enterprises may feel less urgency to adjust their innovation strategies due to "soft budget constraints" and their advantages in resource endowments. In contrast, non-state-owned enterprises face long-term financing constraints and are more sensitive to financial market signals; they must respond swiftly to policy changes to maintain their financing channels and competitive market positions. Therefore, the inhibitory effect of policies on strategic innovation should be more pronounced in non-state-owned enterprises, thereby exerting a stronger dampening effect on green innovation bubbles. Based on this, the following hypotheses are proposed:

H₃: Green finance reform policies may have a stronger dampening effect on green innovation bubbles in non-state-owned enterprises than in state-owned enterprises.

6.2.1 Heterogeneity in Ownership Structure

Heterogeneity in Ownership Structure. Table 7 presents the results of the group-specific tests.

Table 7: Heterogeneity Analysis: Nature of Property Rights

Variable name	TGU	GBR		
	(1) Non-state-owned enterprise	(2) State-owned enterprise	(3) Non-state-owned enterprise	(4) State-owned enterprise
DID	-0.231*** (-4.22)	-0.366*** (-4.35)	-0.0401*** (-9.52)	0.00842 (0.41)
TGU			0.0892*** (15.15)	0.0833*** (18.47)
Control variables	Yes	Yes	Yes	Yes
City fixed effects	Yes	Yes	Yes	Yes
Time fixed effects	Yes	Yes	Yes	Yes
Firm fixed effects	Yes	Yes	Yes	Yes
N	14590	18858	14590	18858
R ²	0.471	0.472	0.467	0.457

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively; the values in parentheses are *t*-statistics adjusted for clustering at the city level. The test for differences in coefficients across groups was based on 500 bootstrap repetitions. In the strategic innovation equation, the difference in DID coefficients across groups was not significant ($p = 0.192$); in the green innovation bubble equation, the difference in DID coefficients across groups was significant ($p = 0.046$).

In terms of strategic innovation, the DID coefficients for non-SOEs and SOEs were -0.231 and -0.366, respectively, both significant at the 1% level. However, a test of the difference in coefficients between the two groups showed no significant difference ($p=0.192$), indicating that there is no significant difference in the inhibitory effect of policy on strategic innovation between the two types of enterprises.

In terms of strategic innovation, the DID coefficients for non-SOEs and SOEs were -0.231 and -0.366, respectively, both significant at the 1% level. However, a test of the difference in coefficients between the two

groups showed no significant difference ($p=0.192$), indicating that there is no significant difference in the inhibitory effect of policy on strategic innovation between the two types of enterprises H3.

In terms of strategic innovation, the DID coefficients for non-SOEs and SOEs were -0.231 and -0.366, respectively, both significant at the 1% level. However, a test of the difference in coefficients between the two groups showed no significant difference ($p=0.192$), indicating that there is no significant difference in the inhibitory effect of policy on strategic innovation between the two types of enterprises.

6.2.2 Regional Heterogeneity

Regarding the regional heterogeneity of green finance policy effects, existing literature has accumulated a wealth of empirical evidence. In terms of the level of green finance development, China exhibits a distinct spatial pattern characterized by “higher levels in the east and lower levels in the west”; specifically, the level of green finance development is relatively higher in the eastern regions, while it lags behind in the central and western regions [15, 16]. In terms of policy impact, the role of green finance in promoting regional development is more pronounced in the eastern regions; for example, its effect on enhancing macroeconomic resilience is stronger in eastern provinces than in central and western provinces. In terms of transmission mechanisms, the higher level of marketization and greater efficiency in resource allocation in the eastern regions are key factors enabling these policy effects to be more fully realized; the carbon reduction effects of green credit are more evident in areas with high resource allocation efficiency, such as eastern cities [17].

Based on this, this paper further examines the regional heterogeneity of the policy effects of green finance reform pilot zones, hypothesizing that the policy’s dampening effect on green innovation bubbles may be stronger in the eastern regions than in the central and western regions. This paper divides the sample into eastern and central-western regions for group-specific regression analysis and employs a permutation test to examine differences in coefficients between groups. Table 8 reports the results of the group-specific regression analysis.

Table 8: Regional heterogeneity: Eastern VS Midwest

Variable name	<i>TGU</i>	<i>GBR</i>		
	(1) East	(2) Mid west	(3) East	(4) Mid west
<i>DID</i>	-0.306***	-0.277***	-0.021***	0.006
	(0.049)	(0.044)	(0.007)	(0.006)
<i>TGU</i>			0.084***	0.090***
			(0.004)	(0.006)
City control effect	Yes	Yes	Yes	Yes
City control effect	Yes	Yes	Yes	Yes
Year control effect	Yes	Yes	Yes	Yes
Firm control effect	Yes	Yes	Yes	Yes
N	24503	9276	24503	9276
R ²	0.473	0.461	0.468	0.463

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively; the values in parentheses are t-statistics adjusted for clustering at the city level. The test for differences in coefficients between groups was based on 500 bootstrap repetitions. The difference in DID coefficients between the two groups in the strategic innovation equation was not significant ($p = 0.444$), and the difference in DID coefficients between the two groups in the green innovation bubble equation was also not significant ($p = 0.218$).

In terms of strategic innovation, the DID coefficients for the eastern and central-western groups were -0.306 and -0.277, respectively, both significant at the 1% level; however, a test of the difference in coefficients between the groups showed no significant difference ($p > 0.1$), indicating that there is no significant difference in the inhibitory effect of the policy on strategic innovation among firms in these two types of regions.

Regarding green innovation bubbles, the DID coefficient for the Eastern group is -0.021, which is significant at the 1% level; whereas the DID coefficient for the Central and Western group is 0.006, which is not significant. The test of differences in coefficients between groups indicates that the difference between the two groups is significant ($p < 0.01$). This suggests that the policy’s dampening effect on green innovation bubbles is significant only in the Eastern region and has not yet manifested in the Central and Western regions. A possible reason lies in the fact that financial markets in the eastern region are more developed and information transmission is more efficient, allowing policy signals to be transmitted more effectively to

corporate behavior; whereas in the central and western regions, the degree of marketization is relatively lower, and the implementation of policy effects may face more obstacles.

To examine regional differences in policy effects, this study divides the sample into two groups—the eastern region and the central and western regions—for regression analysis, and employs a permutation test (500 iterations) to test for differences in coefficients between groups.

Regarding strategic innovation, the DID coefficients for the eastern and central-western groups are -0.306 and -0.277, respectively, both significant at the 1% level; however, the p-value for the between-group difference test is 0.444, indicating no statistically significant difference between the two groups' coefficients. Regarding green innovation bubbles, the DID coefficient for the Eastern group is -0.021 (significant at the 1% level), while that for the Central and Western group is 0.006 (not significant); however, the p-value for the test of differences between groups is 0.218, which is also not significant. Therefore, there is no significant regional difference between the Eastern and Central and Western regions in the policy's dampening effect on green innovation bubbles.

This result indicates that the effects of green finance reform policies are consistent across the eastern and central-western regions; differences in the degree of marketization or the level of financial development among regions do not significantly alter the policy's transmission effects. A possible reason lies in the fact that green finance reform policies, as a form of top-level design, directly impact enterprises through mechanisms such as credit review and information disclosure. Their efficacy can transcend regional barriers, reflecting a certain degree of "universal" characteristics.

7. Conclusions and Policy Recommendations

This study examines Chinese A-share listed companies from 2010 to 2024. Drawing on the quasi-natural experiment created by the establishment of Green Finance Reform and Innovation Pilot Zones, it systematically investigates the impact of green finance policies on corporate green innovation bubbles and their underlying mechanisms. The study finds the following: First, policies in green finance reform pilot zones significantly suppress corporate green innovation bubbles. After policy implementation, the proportion of firms in the treatment group exhibiting bubbles decreased by approximately 3.9 percentage points, equivalent to 35.1% of the mean; this conclusion remains valid after a series of robustness tests. Second, strategic green innovation is the core mechanism through which the policy suppresses bubbles. The policy significantly suppressed firms' strategic green innovation (coefficient = -0.319), whereas strategic innovation significantly exacerbated green innovation bubbles. Path decomposition revealed that the strategic innovation path contributed -0.027, accounting for 69.5% of the policy's total effect, while the direct effect was not significant. Third, policies exert a short-term crowding-out effect on substantive green innovation, but they inherently serve to suppress bubbles. Policies significantly suppressed substantive green innovation (coefficient = -0.129), while substantive innovation significantly suppressed bubbles (coefficient = -0.0046). The contribution of this path was only 0.0006, which was not statistically significant, indicating that the policy crowding-out effect and the innovation-suppressing effect offset each other. Fourth, the policy effects exhibit heterogeneity. In terms of ownership structure, there is no significant difference in the policy's suppression of strategic innovation between the two types of firms, but the suppression of green innovation bubbles is significant only in non-state-owned enterprises (non-SOEs: -0.040; SOEs: 0.008), with a significant difference between groups ($p = 0.034$). In terms of regional heterogeneity, the policy's suppression of strategic innovation was significant in both the eastern and central-western regions, while its suppression of the green innovation bubble was significant only in the eastern region; however, the test for differences between groups was not significant ($p = 0.218$). This indicates that the policy effects exhibit consistency across regions, reflecting the inclusive nature of market-based tools.

This paper identifies strategic green innovation as the primary mechanism through which policies can curb green innovation bubbles; however, whether there are other potential channels (such as corporate green culture and green transmission through supply chains) remains to be explored in future research.

Based on the above conclusions, this paper offers the following policy implications:

First, we should continue to deepen the development of green finance reform pilot zones and strengthen the signal-discrimination function of policies. This paper demonstrates that policies effectively reduce green

innovation bubbles by curbing strategic innovation, indicating that their built-in incentive-compatibility mechanisms—raising signal discrimination standards and altering the relative costs of signaling—play a crucial role. It is recommended that, building on existing pilot programs, the scope of pilot zones be further expanded, successful experiences be promoted, and supporting systems such as environmental information disclosure and green credit review be continuously improved to strengthen the policy’s “correction” function.

Second, establish a differentiated innovation quality assessment system to accurately identify enterprises’ strategic innovation behaviors. Strategic green innovation is a direct source of bubble formation; regulatory authorities should establish a more refined green patent quality assessment system that focuses not only on patent quantity but also on technical content, application scenarios, and environmental benefits. The proportion of green utility model patents could serve as a monitoring indicator, with enterprises showing abnormally high proportions subject to focused scrutiny to reduce the scope for “greenwashing” practices.

Third, address adjustment costs during policy transitions and provide transitional support for substantive innovation. The short-term crowding-out effect of policies on substantive green innovation indicates that enterprises face resource reallocation costs and barriers to technological accumulation during the innovation transition. It is recommended that, during the initial implementation phase of policies, enterprises be provided with a buffer period and transitional support—such as establishing green innovation transition funds and offering technical consulting and talent training services—to help them smoothly shift from strategic innovation to substantive innovation.

Fourth, improve the green credit allocation mechanism to ensure financial resources flow toward genuine green innovation. Although the path of green credit share did not pass the validity test in this study, the core function of green finance remains to direct credit resources toward green sectors. The screening criteria for green credit should be further optimized to prevent credit funds from being used for “greenwashing” projects, ensuring that financial support truly benefits enterprises engaged in high-quality green R&D.

References

- [1] Li, Ruiqing & Han, Jingwang. (2025). The Impact of Green Finance on Urban New-Quality Productivity: A Quasi-Natural Experiment Based on Policies in Green Finance Reform and Innovation Pilot Zones. *Research on Technology, Economy, and Management*, (09), 102–108.
- [2] Cui, Xiao & Tu, Qiangnan. (2025). A Study on the Innovative Incentive Effects of Green Finance Policies: A Quasi-Natural Experiment Based on Green Finance Reform and Innovation Pilot Zones. *China Business Review*, 34(18), 109–113. <https://doi.org/10.19699/j.cnki.issn2096-0298.2025.18.109>.
- [3] Wang Chenyi, Hou Yunkun, Ban Yuanhao & Cheng Yuan. (2025). ESG Ratings and Green Innovation Bubbles: Empirical Evidence from A-Share Listed Companies. *Scientific Decision-Making*, (10), 153–171.
- [4] Tian Lan, Zhao Yufei, & Li Jixiao. (2025). Can Green Finance Policies Empower Enterprises’ New-Quality Productivity Under the “Dual Carbon” Goals?—A Quasi-Natural Experiment Based on Green Finance Reform and Innovation Pilot Zones. *Financial Theory and Practice*, (01), 104–118.
- [5] He Wenbin & Han Xue. (2026). Can Low-Carbon City Pilot Policies Curb Corporate Greenwashing? *Journal of Hefei University of Technology (Social Sciences Edition)*, 40(02), 104–116.
- [6] He Wenbin, Han Xue. Can Low-Carbon City Pilot Policies Curb Corporate Greenwashing? . *Journal of Hefei University of Technology (Social Sciences Edition)*, 2026, 40(02): 104–116.
- [7] Shen Yu, Huang Hao, & Zhao Ling. (2018). Local Governments’ “Innovation Worship” and Corporate Patent Bubbles. *Science and Technology Management*, 39(04), 83–91. <https://doi.org/10.19571/j.cnki.1000-2995.2018.04.009>.
- [8] Yang, G. C., & Rui, M. (2020). Incentive and Compliance Effects of Tax Incentive Policies for High-Tech Enterprises. *Economic Research Journal*, 55(09), 174–191.

- [9] Liu, Jing. (2024). Green Finance, Environmental Regulation, and Strategic Green Innovation. *Journal of Dalian University of Technology (Social Sciences Edition)*, 45(04), 50-59. <https://doi.org/10.19525/j.issn1008-407x.2024.04.006>.
- [10] Luo Shuangcheng, Liu Jianjiang, & Xiong Zhiqiao. (2024). Talent Policy Support and Green Innovation Performance of Heavy Polluters: Empirical Evidence from High-Level Talent Subsidies. *Industrial Economics Research*, (01), 56–70. <https://doi.org/10.13269/j.cnki.ier.2024.01.010>.
- [11] Lu, Y., Tao, Z., & Zhu, L. (2017). Identifying FDI spillovers. *Journal of International Economics*, 107, 75–90.
- [12] Shi Lei, Yang Zhen, & Qian Guiming. (2025). Digital Industrial Cluster Policies and Breakthrough Innovation in Key Core Technologies. *China Industrial Economics*, (01), 100–117. <https://doi.org/10.19581/j.cnki.ciejournal.2025.01.006>.
- [13] Zhou, X., & Yamamoto, T. (2023). Tracing causal paths from experimental and observational data. *The Journal of Politics*, 85(1), 250–265.
- [14] Fang Junxiong. (2007). Ownership Structure, Marketization Process, and Capital Allocation Efficiency. *Management World*, (11), 27-35. <https://doi.org/10.19744/j.cnki.11-1235/f.2007.11.004>.
- [15] He Guosheng & Liu Rongbing. (2024). A Study on Regional Differences and Influencing Factors in the Development of Green Finance in China. *Journal of Guangxi University (Philosophy and Social Sciences Edition)*, 46(05), 100–112. <https://doi.org/10.13624/j.cnki.jgupss.2024.05.030>.
- [16] Xing Yu & Bian Weijun. (2023). Regional Differences and Dynamic Evolution Trends in the Development Efficiency of China's Green Finance. *Xinjiang Social Sciences*, (02), 62–72. <https://doi.org/10.20003/j.cnki.xjshkx.2023.02.007>.
- [17] Shen Lumin, Cai Qian, Yang Yongliang & Liao Zhongju. (2026). The Impact of Green Credit on Urban Carbon Emissions Reduction from a Resource Allocation Perspective. *Science, Technology, and Economy*, 39(01), 106–110. <https://doi.org/10.14059/j.cnki.cn32-1276n.2026.01.022>.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Research on the Impact of Digital Financial Literacy on Household Financial Portfolio Allocation Efficiency

Li Yang^{1,2}, Chen Yang^{1,*}

¹Guangxi Normal University, Guilin, Guangxi, China

²Key Laboratory of Digital Empowerment Economic Development (Guangxi Normal University), Education Department of Guangxi Zhuang Autonomous Region, Guilin, Guangxi, China

*Corresponding author: Chen Yang.

Abstract

Against the backdrop of the burgeoning digital economy, this paper investigates the impact of residents' digital financial literacy on the efficiency of household financial asset allocation and its underlying mechanisms. Using micro-survey data from the China Household Finance Survey (CHFS) for 2017 and 2019, we conduct an empirical test. The findings reveal that an improvement in digital financial literacy significantly enhances the efficiency of household financial asset allocation. This effect is achieved by increasing household financial accessibility and promoting social interaction. Moreover, these results remain robust and significant after replacing the measurement indicators of the explanatory variable and accounting for potential policy interference. The heterogeneity analysis indicates that the positive effect of digital financial literacy is larger for rural households, those in non-key urban agglomerations, and non-financially vulnerable households. This study holds significant implications for enhancing the digital financial literacy of residents in China and promoting the rationalization of household financial asset allocation.

Keywords

digital financial literacy, household financial portfolio allocation efficiency, financial accessibility, social interaction

1. Introduction

According to data from the National Bureau of Statistics of China, in the first quarter of 2025, the national per capita disposable income reached 12179 yuan, with a year-on-year nominal growth of 5.5% and an actual growth rate of 5.6%. Among this, the per capita net property income was 1015 yuan, with a growth of 2.7%, accounting for 8.3% of the disposable income. This demonstrates that China's per capita property income exhibits the characteristics of "slow growth and low proportion", reflecting the relatively limited channels for residents to increase their wealth. As an important way to increase property income, the efficiency of household financial asset allocation is crucial for the preservation and appreciation of wealth.

Digital financial literacy originates from the intersection of financial literacy and digital literacy, with its theoretical connotations continuously enriched by the digital transformation of finance. It has roughly undergone three evolutionary stages: initially focusing on the skill dimension of conducting online financial

activities using digital tools [1]; subsequently emphasizing the comprehensive application of financial knowledge and digital financial skills [2], measuring aspects such as financial knowledge, digital payments, digital wealth management, and digital lending; in recent years, in response to the frequent occurrence of online fraud, scholars have further expanded it to include multi-dimensional abilities covering digital risk prevention and consumer rights protection awareness [3]. Overall, as human capital in the digital age, digital financial literacy reflects the ability of households to make decisions and manage risks using digital tools and financial knowledge. However, the “2023 Global Perspective on China’s Digital Financial Literacy Research Report” indicates that the overall level of digital financial literacy among Chinese residents is low, especially among the elderly aged 60 and above and the youth aged 16–24, who exhibit significant shortcomings. This lack of literacy is likely to be a key constraint hindering families from optimizing asset allocation through digital financial services.

The ultimate objective of household financial asset investment is to maximize portfolio returns per unit of risk. Accordingly, some scholars have measured the efficiency of household financial asset allocation based on a comprehensive integration of risk and return [4]. A multitude of factors influence this allocation efficiency, encompassing individual cognition, household characteristics, background risks, and the external economic environment [5–7]. Existing literature demonstrates that enhancements in residents’ financial literacy, digital literacy, and the usage of digital finance can effectively optimize household financial behaviors—such as consumption, saving, and investment—thereby significantly improving household financial well-being [8,9]. These findings provide a robust theoretical foundation for research concerning the impact of digital financial literacy on household financial behavior.

Based on this, the paper utilizes data from the 2017 and 2019 rounds of the China Household Finance Survey to explore the relationship between digital financial literacy and the efficiency of Chinese household financial asset allocation, and further analyzes its impact mechanism. The marginal contributions of this paper are reflected in the following aspects: Firstly, innovation in research perspective. First, this study offers an innovative research perspective. On the one hand, by deconstructing digital financial literacy into three dimensions—knowledge, skills, and risk awareness—it accurately characterizes residents’ financial behavioral capabilities in the digital era. On the other hand, whereas the traditional breadth and depth of financial market participation merely reflect portfolio risk or investment scale, this paper comprehensively integrates both portfolio risk and return to measure the efficiency of household financial asset allocation. Secondly, innovation in research mechanism. Few scholars have directly studied the impact of digital financial literacy on the efficiency of household financial asset allocation, lacking exploration of its intrinsic action path and transmission mechanism. This paper explores the specific mechanism of the impact of residents’ digital financial literacy on the efficiency of household financial asset allocation from the perspective of household financial accessibility and social interaction.

The subsequent parts of this paper are arranged as follows: Part II presents theoretical analysis and hypothesis formulation; Part III introduces the data, variables, and model; Part IV analyzes the empirical results of the impact of digital financial literacy on the efficiency of household financial asset allocation; and Part V concludes the research and offers policy suggestions.

2. Research Hypotheses

2.1 Digital Financial Literacy and Household Financial Portfolio Allocation Efficiency

In reality, due to cognitive limitations, information asymmetry, and subjective preferences, household asset allocation often deviates from the efficient frontier. As a key human capital in the digital age, the enhancement of digital financial literacy is mainly reflected in the improvement of financial knowledge, the advancement of digital financial skills, and the strengthening of digital financial risk awareness. These three aspects can effectively mitigate behavioral biases of investors [10], prompting household investment portfolios to approach the efficient boundary, and thereby enhancing risk-adjusted returns, which is to improve the efficiency of household financial asset allocation. The specific mechanisms are as follows:

Firstly, the improvement of residents’ financial knowledge level can correct cognitive biases and enhance information processing capabilities, thus optimizing household financial asset portfolios. Investors with a high level of financial knowledge are better equipped to accurately understand market information and product

characteristics, thereby correcting cognitive biases caused by knowledge deficiency and reducing irrational investment. Facing risk assets with high information density, solid financial knowledge can significantly shorten the time for information acquisition and understanding, reducing the information costs of entering risk markets [11]. Additionally, digital financial literacy enhances households' sensitivity to market fluctuations, enabling them to dynamically adjust their asset portfolio structures in a more timely manner [9], thereby optimizing the efficiency of household financial asset allocation.

Secondly, the advancement of digital financial skills can reduce market frictions and promote asset portfolio diversification, thus optimizing household financial asset portfolios. Proficient digital skills enable households to access online platforms more conveniently, significantly reducing the time and capital costs of information acquisition and transaction execution [9,12]. This not only enhances households' willingness to participate in risk markets [11] but also broadens product selection channels, facilitating the transformation of household investment portfolios from single to diverse, effectively diversifying non-systematic risks [13]. Moreover, leveraging the advantages of big data and algorithms of financial institutions, households with strong digital financial skills can better utilize digital tools to accurately match their needs with risk-appropriate products, thereby reducing asset misallocation risks and improving asset allocation efficiency [14].

Thirdly, the enhancement of digital financial risk awareness can correct the deviation between subjective risk preferences and objective risk-taking capacity, thereby improving the efficiency of household financial asset allocation. Insufficient risk perception often leads households to overestimate market risks, causing their subjective risk preferences to be far below their actual risk tolerance levels, which in turn results in excessive concentration of assets in low-yield risk-free assets, missing out on risk premiums. However, the increase in risk awareness encourages households to rationally assess their own risk tolerance and product risk characteristics [15]. This not only enhances the willingness to participate in the market but also guides households to avoid excessive conservativeness within acceptable risk levels, obtaining risk premiums through reasonable allocation of risk assets, ultimately improving the overall efficiency of the asset portfolio [9].

Hypothesis 1. *Enhancing residents' digital financial literacy positively affects the efficiency of household financial asset allocation.*

2.2 Impact Mechanism of Digital Financial Literacy on Household Financial Portfolio Allocation Efficiency

2.2.1 Financial Accessibility

Financial accessibility refers to the ability of residents to equitably and conveniently access and utilize formal financial products and services at an affordable cost, overcoming geographical, informational, and other barriers. Residents' digital financial literacy—encompassing the comprehension of financial knowledge, the application of digital financial tools, and the awareness of digital risks—directly determines their capacity to effectively acquire and utilize financial resources. Specifically, digital financial literacy enhances household financial accessibility from the following three dimensions, thereby optimizing asset allocation efficiency:

First, it lowers transaction costs and expands the asset choice set, thereby enhancing the accessibility of financial products. Traditional financial services are associated with high information and transaction costs, which have historically led to low household market participation. Although digital finance has mitigated these frictions, it has concurrently introduced risks associated with identifying informal financial products. Residents with high digital literacy can proficiently navigate online platforms to effectively evade informal assets and digital fraud. Consequently, they can access diversified formal financial products, such as equities and mutual funds, at a lower cost [16], thereby strengthening their willingness to participate in risk-bearing markets [11]. According to Modern Portfolio Theory, an extensive asset choice set is a prerequisite for constructing an efficiently diversified portfolio. Therefore, digital literacy empowers households to allocate assets within a safer and broader scope, optimizing the risk-return profile of their investment portfolios through risk diversification mechanisms.

Second, it facilitates households' utilization of professional financial services, thereby improving the accessibility of financial services. Digitalization has given rise to innovative services such as robo-advisors. However, effectively leveraging these services requires overcoming the “digital threshold” and mitigating “self-exclusion” stemming from a lack of knowledge [17]. A higher level of digital literacy not only alleviates

this self-exclusion [13] but also helps households overcome technological barriers, increasing their willingness to proactively seek professional financial advice via digital channels [18]. Such professional services can compensate for households' inherent knowledge deficits, guiding them toward scientific and personalized asset allocation and directly enhancing allocation efficiency.

Finally, it strengthens the capacity for information acquisition and discernment, thereby improving the accessibility of financial information. A high level of digital financial literacy endows households with the ability to efficiently acquire and deeply process financial information. This enables households to make rational decisions based on a more comprehensive information set and avoid herd behavior. Consequently, it reduces the risk of asset misallocation caused by information asymmetry and significantly improves the rationality of their financial decision-making. Taken together, these mechanisms suggest a mediating role for financial accessibility.

Hypothesis 2. *The enhancement of digital financial literacy improves the efficiency of household financial asset allocation through the channel of increased household financial accessibility.*

2.2.2 Social Interaction

Social interaction is defined as the process in which individuals initiate social actions toward others and receive reactive responses. In the context of household economic behavior, social interaction emphasizes that residents' participation in financial markets is influenced by other members of their social groups. The enhancement of digital financial literacy drives the extension of households' social interactions and economic transactions into virtual networks, increasing the frequency of social interactions by lowering interaction costs. This not only consolidates existing kinship-based social networks but also expands Internet-based social circles, thereby enhancing the efficiency of household financial asset allocation by optimizing the information environment.

On the one hand, social interaction alleviates the constraints of bounded rationality by broadening information sources. Digital interaction platforms transcend the geographical boundaries of traditional social networks [2], thereby improving the efficiency of information exchange [19]. According to bounded rationality theory, households often struggle to make optimal decisions due to insufficient cognitive and information-processing capacities. Residents with high digital financial literacy can proactively utilize online platforms to acquire supplementary information, such as market trends [8], effectively mitigating asset misallocation caused by information asymmetry and lags [20]. Furthermore, leveraging their financial knowledge reserves, they can conduct in-depth analyses of the information gathered through interactions [21], providing a reliable basis for investment decisions and consequently improving portfolio returns.

On the other hand, social interaction may also trigger collective irrational behaviors, such as herd behavior and trend-chasing (buying high and selling low). However, households with high digital financial literacy can leverage their cognitive reserves to rationally screen and analyze the financial information acquired through social interactions. This enables them to overcome ambiguity aversion induced by information asymmetry and curb the aggressive sentiments associated with overconfidence. This process effectively averts decision-making errors driven by emotional biases, prompting households to construct more rational financial asset portfolios [22].

Hypothesis 3. *The enhancement of digital financial literacy improves the efficiency of household financial asset allocation through the channel of facilitating household social interaction.*

3. Research Design

3.1 Data and Sample

Considering that the COVID-19 pandemic, as a major public health emergency, has inflicted severe exogenous shocks on both the macroeconomic environment and micro-level household decision-making, this paper selects data from the 2017 and 2019 waves of the China Household Finance Survey (CHFS). To ensure data quality, the following preprocessing procedures are implemented:

(1) Considering the minimum age restrictions for accessing financial assets and the constraint that cognitive decline among the advanced elderly leads to intergenerational transfers in household financial decision-making,

only samples with household heads aged between 18 and 85 are retained.(2) Samples with negative total household income or negative total household assets are excluded.(3) Observations with missing values or outliers in the relevant variables are dropped.(4) Continuous variables, such as total assets and total income, are log-transformed and subsequently winsorized at the 1st and 99th percentiles to mitigate estimation biases caused by extreme sample outliers.

Ultimately, an unbalanced panel dataset spanning two waves (2017 and 2019) is constructed, yielding a total of 25,801 valid observations.

3.2 Variables and Definitions

3.2.1 Dependent Variable

The dependent variable is the efficiency of household financial asset allocation. We employ $Sharp\_Ratio_{it}$ and $Sortino\_Ratio_{it}$ as proxy variables for this efficiency [6]. The specific construction methods are as follows:

First, household financial assets are categorized into risk-free and risky financial assets. Subsequently, risky financial assets are classified into three major types—bonds, funds, and equities—based on their risk-return profiles. Given that the CHFS dataset only reports the amounts of financial assets allocated by households rather than the specific returns of each asset type, this paper employs market indices as proxies for the returns on various financial assets. Subsequently, the average return method is applied to measure the annual returns of different financial assets for each household [5]. Specifically, the proxy indices for each type of financial asset are determined as shown in Table 1. The historical average returns from January 2003 to December 2017 are utilized to represent the household's returns on various financial assets for the year 2017; the same procedure is applied for the year 2019.

Table 1: Alternative Indices for Financial Asset Returns and Data Sources.

Return Type	Proxy Method	Data Source
Equity asset return	Turnover-weighted average return of the Shanghai Composite Index and the Shenzhen Component Index	China Economic and Financial Database
Bond asset return	Return on the CSI Aggregate Bond Index	Wind Database
Fund asset return	Turnover-weighted average return of the Shanghai Fund Index and the Shenzhen LOF Fund Index	Wind Database
Risk-free rate	One-year lump-sum time deposit rate	People's Bank of China

From the above data, we derive the time series of excess returns and volatilities for various financial assets, which allows us to calculate the Sharpe and Sortino ratios for the household financial asset portfolio. The specific formulas are as follows:

$$Sharp\_Ratio_{it} = [E(R_{i,t}) - R_{f,t}] / \delta_{i,t} \quad (1)$$

$$Sortino\_Ratio_{it} = [E(R_{i,t}) - R_{f,t}] / \delta_{downi,t} \quad (2)$$

$$E(R_{i,t}) = \sum_{j=1}^N W_{ji,t} R_{ji,t} \quad (3)$$

$$\delta_{i,t}^2 = \sum_{j=1}^N \sum_{k=1}^N W_{ji,t} W_{ki,t} \sigma(R_{ji,t}, R_{ki,t}) \quad (4)$$

where $Sharp\_Ratio_{it}$ and $Sortino\_Ratio_{it}$ denote the Sharpe ratio and Sortino ratio of household i 's financial asset portfolio in period t . $E(R_{i,t})$ and $\delta_{i,t}$ represent the expected return and standard deviation of the portfolio in period t . $R_{f,t}$ is the risk-free rate. $R_{ji,t}$ represents the return on asset class j held by household i in period t . $W_{ji,t}$ denotes the weight of asset class j in the total financial assets of household i in period t . $\sigma(\cdot)$ denotes the covariance. $\delta_{downi,t}$ is the downside standard deviation, in the Sortino ratio only accounts for the volatility of returns falling below the risk-free rate.

3.2.2 Key Explanatory Variable

The core explanatory variable in this study is DFL. First, we measure respondents' financial knowledge using interest rate calculation and inflation calculation [1]. Meanwhile, we construct two dummy variables for each of these questions: one indicating whether the answer is correct, and the other indicating whether the respondent provided a direct answer [8]. Second, to measure individuals' ability to apply digital financial tools in practice, we use several indicators, including whether the respondent holds yield-generating internet wealth management products, borrows through internet platforms, uses mobile payment, and holds stock accounts, mutual fund accounts, or credit cards. Finally, the dimension of digital financial risk awareness is designed to capture individuals' perception and understanding of the specific risks associated with digital financial products. Scores are assigned based on two items: the "comparison of risks between stocks and funds" and the "subjective judgment of the risks associated with internet wealth management products".

3.2.3 Mechanism Variables

First, financial accessibility. Based on the definition of financial accessibility, an index is constructed covering three dimensions: the accessibility of financial products, financial services, and financial information [13,23]. The specific indicators selected are as follows: (1) The number of bank accounts per capita in the household. (2) Whether the household's credit demand is satisfied. A value of 1 is assigned if the credit demand is met, and 0 if the demand exists but remains unmet. (3) The number of accounts opened by the household for purchasing risky financial assets. (4) Whether the household has access to financial or investment advisors. (5) The degree of the household's attention to financial information. A value of 2 is assigned for "very or fairly concerned", 1 for "average or rarely concerned", and 0 for "never concerned". (6) Whether the household obtains financial information via the Internet or wealth management apps (assigned a value of 1 for "yes" and 0 otherwise). Finally, PCA is employed on the above indicators to construct a comprehensive index of household financial accessibility.

Second, social interaction. We employ Principal Component Analysis (PCA) to construct a comprehensive index of social interaction, utilizing the logarithms of expenditures on weddings and funerals, holiday cash gifts, transportation, and communication [20].

3.2.4 Control Variables

This study selects control variables at three levels: the household head, the household, and the regional economy. Specifically, these include:

(1) Household head level: age of the household head and its square, gender (1 = male, 0 = female), marital status (1 = married, 0 otherwise), education level (an ordinal variable ranging from 1 to 6, where 1 = no schooling, 2 = primary school, 3 = junior high school, 4 = senior high school/vocational school, 5 = junior college/undergraduate, and 6 = postgraduate or above), registration type (4 = urban, 3 = unified, 2 = rural, 1 = other), and social pension insurance participation (1 if the household head is covered, 0 otherwise).

(2) Household level: old-age dependency ratio (the proportion of household members aged 60 and above), child dependency ratio (the proportion of members aged 16 and below), the logarithms of total household income and total household assets, and the share of non-financial assets.

(3) Regional economic level: economic development level (the logarithm of per capita GDP by province), regional population density, and the level of digital financial inclusion.

3.3 Model Specification

To address the sample selection problem, we employ the Heckman two-step procedure for estimation [5]. In the first step, a Probit model is used to estimate the selection equation for household participation in risky financial markets. The specific model is specified as follows:

$$riskfin_{it} = \begin{cases} 1, & \text{若 } riskfin_{it}^* > 0 \\ 0, & \text{若 } riskfin_{it}^* \leq 0 \end{cases} \quad (5)$$

$$riskfin_{it}^* = \gamma_0 + \gamma_1 DFL_{it} + \gamma_2 X_{it} + \gamma_3 unhealthy_{it} + Province_i + Year_t + u_{it} \quad (6)$$

Equation (5) is the equation for the binary variable $riskfin_{it}$. The value of $riskfin_{it}$ depends on an unobservable latent variable $riskfin_{it}^*$. If $riskfin_{it}^* > 0$, indicating that the household participates in the risky financial market, then $riskfin_{it} = 1$; If $riskfin_{it}^* \leq 0$, indicating that the household does not participate in the risky financial market, then $riskfin_{it} = 0$. i and t represent the observed household and year. u_{it} is distributed as normal. In Equation (6), DFL_{it} denotes the residents' digital financial literacy. X_{it} denotes the control variables. $Province_i$ and $Year_t$ denote province and time fixed effects, respectively. In addition, an exclusion restriction needs to be added to the selection equation in the first step. Since family members' health status significantly affects household participation in risky financial markets yet exerts no direct effect on the efficiency of financial asset portfolios [9], we use $unhealthy_{it}$ to serve as the exclusion restriction.

Next, the Inverse Mills Ratio λ_i is derived from the estimation results of the first-stage selection equation. The calculation formula is given by Equation (7). $\varnothing(\cdot)$ is the probability density function of the standard normal distribution. $\Phi(\cdot)$ is the standard normal cumulative distribution function. W_i' is the vector of key explanatory and control variables. $\hat{\gamma}$ is estimated coefficient vector.

$$\lambda_i = \varnothing(-W_i' \hat{\gamma}) / \Phi(-W_i' \hat{\gamma}) \quad (7)$$

In the second step, we estimate the impact of digital financial literacy on household financial asset allocation efficiency using OLS, and incorporate λ_i as an explanatory variable into Equations (8) and (9).

$$Sharp\_Ratio_{it} = \alpha_0 + \beta_0 DFL_{it} + \delta_0 X_{it} + \vartheta_0 \lambda_i + Province_i + Year_t + \varepsilon_{it} \quad (8)$$

$$Sortino\_Ratio_{it} = \alpha_1 + \beta_1 DFL_{it} + \delta_1 X_{it} + \vartheta_1 \lambda_i + Province_i + Year_t + \varepsilon_{it} \quad (9)$$

In Equations (8) and (9), $Sharp\_Ratio_{it}$ and $Sortino\_Ratio_{it}$ denote the Sharpe ratio and the Sortino ratio of the household financial asset portfolio, respectively, with higher values indicating greater efficiency in household financial asset allocation. If the coefficients of the Inverse Mills Ratio ϑ_0 and ϑ_1 are significantly non-zero, indicating the presence of sample selection bias, the Heckman two-step method is required to obtain unbiased estimation results. ε_{it} is the random disturbance term. The other variables are defined as in Equation (2).

Additionally, to test Hypotheses 2 and 3, two mechanism variables reflecting the impact of digital financial literacy on household financial asset allocation efficiency are proposed based on economic theory. Then, an OLS model is employed to identify the causal effect of digital financial literacy on the mechanism variables. The specific model specification is as follows:

$$M_{it} = \alpha_2 + \beta_2 DFL_{it} + \delta_2 X_{it} + Province_i + Year_t + \varepsilon_{it} \quad (10)$$

In Equation (10), M_{it} represents the mechanism variables, denoting financial accessibility (ss) and social interaction ($Social\_Interaction_{it}$), respectively. The other variables are defined as in Equations (8) and (9).

3.4 Descriptive Statistics

Table 2 reports the summary statistics of the variables. A multicollinearity test shows that all VIF values are below 10, ruling out any multicollinearity concerns and confirming the soundness of the selected variables.

Table 2: Results of descriptive statistics.

Variable	Observations	Mean	SD	Min	Max
DFL	25801	1.032	0.616	0.000	2.938
Sharpe	25801	0.431	0.262	0.000	0.984
Sortino	25801	2.322	2.131	0.000	14.571
age	25801	52.564	14.169	18.000	85.000
Age_squared	25801	29.638	15.121	3.240	72.250
Gender	25801	0.769	0.421	0.000	1.000
Marriage	25801	0.874	0.332	0.000	1.000
Edu	25801	3.630	1.131	1.000	6.000

Egistration	25801	3.035	0.910	1.000	4.000
Pension	25801	0.854	0.353	0.000	1.000
Older	25801	0.279	0.386	0.000	1.000
Childer	25801	0.118	0.167	0.000	0.778
Total_income	25801	11.283	1.162	5.775	13.477
Total_asset	25801	13.629	1.384	7.959	16.178
Nonfin	25801	0.755	0.240	0.000	1.000
LnGDP	25801	11.129	0.413	10.279	11.994
Pop_density	25801	5.982	1.055	2.248	7.996
DIFI	25801	5.722	0.127	5.500	6.019

Notes: Mean, Std. Dev., Min, and Max are rounded to three decimal places.

4. Empirical Analysis

4.1 Baseline Regression

Table 3 reports the estimation results regarding the impact of digital financial literacy on the efficiency of household financial asset allocation. The first-stage results of the Heckman model in Column (1) reveal that the marginal effect of digital financial literacy on participation in risky financial markets is 0.2522, which is significant at the 1% level. This indicates that households with higher levels of digital financial literacy exhibit a stronger willingness to participate in risky financial markets. Furthermore, the inverse Mills ratio in the second stage is positive and statistically significant, confirming that OLS estimates suffer from sample selection bias and thereby validating the appropriateness of employing the Heckman model.

The second-stage results in Columns (2) and (3) demonstrate that the coefficients of digital financial literacy on the Sharpe ratio and the Sortino ratio are 0.1499 and 0.7701, respectively, both significant at the 1% level. This suggests that digital financial literacy significantly enhances the excess returns per unit of total risk and per unit of downside risk borne by households, thereby substantially improving asset allocation efficiency. Notably, its enhancing effect is more pronounced on the Sortino ratio, indicating that a high level of digital literacy significantly strengthens households' ability to manage downside risk, consequently bolstering the robustness of their portfolios against market volatility. In summary, Hypothesis 1 is supported.

Table 3: Baseline Regression Results.

Variable	(1)	(2)	(3)
	First Stage	Second Stage	Second Stage
	Risky Financial Market Participation	Sharpe	Sortino
DFL	0.2522 *** (0.0035)	0.1499 *** (0.0049)	0.7701 *** (0.0536)
Unhealthy	-0.0207 *** (0.0069)		
IMR		0.0320 *** (0.0105)	-0.1712 * (0.1022)
Control Variables	YES	YES	YES
Year Fixed Effects	YES	YES	YES
Province Fixed Effects	YES	YES	YES
N	54675	25801	25801
Pseudo R^2/R^2	0.3188	0.2514	0.1002

Notes: (1) ***, **, * denote significance at the 1%, 5%, and 10% levels, respectively. This notation applies hereafter. (2) Robust standard errors are reported in parentheses. The first column reports marginal effects.

There may be endogeneity issues between digital financial literacy and the efficiency of household financial asset allocation. On the one hand, the experience households accumulate during investment practices can enhance their digital financial literacy, and this reverse causality may lead to an overestimation of the results. On the other hand, unobservable omitted variables may also cause estimation bias. To address these issues,

this paper selects “the average digital financial literacy of other households in the community” as an instrumental variable.

Table 4 reports the IV regression results. The first-stage results show that the IV is significantly and positively correlated with digital financial literacy. The Kleibergen-Paap rk LM statistic is 836.98, which is significant at the 1% statistical level, and the Cragg-Donald Wald F statistic is 999.78, exceeding the critical value of 16.38 at the 10% level. These results rule out the problems of underidentification and weak instruments. The second-stage results demonstrate that the coefficients of digital financial literacy on the Sharpe ratio and the Sortino ratio remain significantly positive, which is highly consistent with the baseline regression, thereby further supporting Hypothesis 1.

Table 4: *Instrumental Variable Regression Results.*

Variable	(1) Sharpe	(2) Sortino
DFL	0.0565 *** (0.0148)	0.5317 *** (0.1317)
Control Variables	YES	YES
Year Fixed Effects	YES	YES
Province Fixed Effects	YES	YES
N	25801	25801
First Stage		
Community average household digital financial literacy	0.4122 *** (0.0135)	
Kleibergen-Paap rk LM	836.98 ***	
Cragg-Donald Wald F	999.78	

4.2 Robustness Test

4.2.1 Replacing the Core Explanatory Variable

To test the robustness of the results, this paper remeasures digital financial literacy using the entropy method and the score summation method. The regression results based on the entropy method in Columns (1) and (2) of Table 5 show that the coefficients of digital financial literacy on the Sharpe ratio and the Sortino ratio of household financial portfolios are 0.3906 and 3.9965, respectively, both significant at the 1% level. Meanwhile, the results based on the score summation method in Columns (3) and (4) yield corresponding coefficients of 0.0369 and 0.1293, which are similarly significant at the 1% level. The above findings are highly consistent with the baseline regression, reaffirming that an increase in digital financial literacy helps enhance the risk-adjusted returns of household financial portfolios—thereby improving the efficiency of household financial asset allocation. Thus, the baseline regression results remain robust.

Table 5: *Results of Robustness Test by Replacing the Explanatory Variable.*

Variable	Entropy Weight Method		Equal-Weighted Summation Method	
	(1)	(2)	(3)	(4)
	Sharpe	Sortino	Sharpe	Sortino
DFL	0.3906 *** (0.0117)	3.9965 *** (0.1544)	0.0369 *** (0.0012)	0.1293 *** (0.0134)
IMR	−0.1526 *** (0.0066)	−0.7506 *** (0.0508)	0.0418 *** (0.0108)	−0.556 *** (0.1133)
Control Variables	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Province Fixed Effects	YES	YES	YES	YES
N	25801	25801	25801	25801
adj. R ²	0.2584	0.1391	0.2498	0.093

4.2.2 Excluding Policy-Interfered Samples

The “Broadband China” strategy and “Smart City” pilots have continuously improved China’s digital infrastructure and economic ecosystem, prompting residents to significantly enhance their digital financial literacy through “learning by doing” during high-frequency digital interactions. This leads to significantly higher literacy levels among residents in pilot cities. To exclude such policy interference, this paper constructs a regional digital infrastructure index based on data from the National Bureau of Statistics, encompassing five dimensions: telephone penetration rate, length of long-distance optical cables, internet penetration rate, number of broadband access ports, and number of domain names [24]. Subsequently, we exclude household samples located in regions with digital infrastructure levels above the average and re-estimate the model. The results in Table 6 show that the coefficient of digital financial literacy on allocation efficiency remains significantly positive, indicating that after controlling for the confounding effects of policy dividends, the baseline regression conclusions remain highly robust.

Table 6: Robustness Check Results Excluding Policy-Interfered Samples.

Variable	(1)	(2)
	Sharpe	Sortino
DFL	0.1749 *** (0.0078)	1.0035 *** (0.0902)
IMR	0.086 *** (0.0154)	0.3948 ** (0.1627)
Control Variables	YES	YES
Year Fixed Effects	YES	YES
Province Fixed Effects	YES	YES
N	13839	13839
adj. R ²	0.2476	0.1214

4.3 Mechanism Tests

To verify whether the enhancement of digital financial literacy improves household financial asset allocation efficiency by increasing household financial accessibility and promoting social interaction, this paper conducts mechanism tests.

Column (3) of Table 7 shows that the coefficient of digital financial literacy on household financial accessibility is 0.536 and significant at the 1% level, indicating that an improvement in residents’ digital financial literacy effectively promotes household financial accessibility. Existing literature indicates that financial accessibility is crucial for enhancing the efficiency of household financial asset allocation, primarily achieved through two pathways: First, it broadens the investment scope, shifting assets from single savings to diversified portfolios such as wealth management products, mutual funds, and stocks, thereby effectively diversifying market risks [16]. Second, it lowers transaction and information costs, reduces market frictions, and incentivizes households to seek professional consultation to overcome cognitive limitations. Together, these two pathways enhance households’ responsiveness to portfolio adjustments, ultimately improving asset allocation efficiency [9]. In summary, Hypothesis 2 is supported.

Column (4) of Table 7 reveals that the coefficient of digital financial literacy on household social interaction is 0.123 and significant at the 1% level, demonstrating that an enhancement in digital financial literacy strengthens households’ participation in social interactions. Previous studies have pointed out that households’ participation in social interactions contributes to improving their financial asset allocation efficiency [20]. Specifically: First, social interactions expand households’ decision-making information sets through information transmission mechanisms, thereby reducing decision-making biases caused by information scarcity [25]. Second, households with high digital financial literacy can leverage their cognitive reserves to rationally filter financial information acquired during interactions, effectively avoiding emotional decision-making errors and facilitating the construction of optimal asset portfolios [22]. In summary, Hypothesis 3 is supported.

Table 7: Mechanism Test Results.

Variable	(1)	(2)	(3)	(4)
	Sharpe	Sortino	Financial_Accessibility	Social_Interaction

DFL	0.1499 *** (0.0049)	0.7701 *** (0.0536)	0.6267 *** (0.009)	0.1525 (0.0072)
Control Variables	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Province Fixed Effects	YES	YES	YES	YES
N	25801	25801	25801	25801
adj. R^2	0.2514	0.1002	0.4257	0.7967

4.4 Heterogeneity Analysis

4.4.1 Urban-Rural Heterogeneity Analysis

Given the unbalanced urban-rural development in China, which leads to significant disparities in digital infrastructure, access conditions, and financial receptivity, this paper divides the sample into urban and rural subsamples to examine the heterogeneous effects of digital financial literacy on household financial asset allocation efficiency.

The subsample regression results in Table 8 indicate that digital financial literacy significantly improves the asset allocation efficiency of both urban and rural households, but this enhancing effect is more pronounced among rural households. A Chow test confirms that the difference between the two groups is significant at the 1% level. This disparity primarily stems from the fact that urban households typically possess extensive traditional financial channels and a solid cognitive foundation; thus, the enhancement of digital financial literacy serves more as a refinement of their existing knowledge system, yielding relatively limited marginal improvements. In contrast, rural households have long been constrained by insufficient financial accessibility and information isolation. For them, improved digital financial literacy effectively helps overcome information and channel barriers, substantially optimizing their asset allocation behavior and thereby exhibiting higher marginal effects.

Table 8: Regression Results for Urban-Rural Heterogeneity.

Variable	Sharpe		Sortino	
	(1)	(2)	(3)	(4)
	Urban	Rural	Urban	Rural
DFL	0.1364 *** (0.0055)	0.1742 *** (0.0162)	0.7016 *** (0.0589)	0.8171 *** (0.1278)
Control Variables	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Province Fixed Effects	YES	YES	YES	YES
N	21036	4765	21036	4765
adj. R^2	0.2502	0.3044	0.0798	0.3818
Chow test P-value	0.0000		0.0000	

4.4.2 Regional Heterogeneity Analysis

According to China's regional development planning, this paper divides the sample into key and non-key urban agglomerations to examine regional heterogeneity. The subsample regression results in Table 9 show that digital financial literacy has a significant positive impact on the asset allocation efficiency of both groups of households, but this enhancing effect is more pronounced in non-key urban agglomerations. A Chow test confirms that the cross-group difference is significant at the 1% level.

The underlying logic of this phenomenon is twofold. On the one hand, for households in non-key regions, improvements in digital financial literacy enable them to effectively leverage digital tools to overcome geographical constraints and access a broader range of financial resources, thereby substantially improving their asset allocation efficiency. On the other hand, the high living costs in key urban agglomerations—particularly regarding housing, education, and healthcare—exert a significant crowding-out effect on households' investable funds. To meet daily expenditures, these households exhibit a preference for highly

liquid and low-risk assets, which consequently weakens the marginal effect of enhanced digital financial literacy.

Table 9: Regional Heterogeneity Test Results.

Variable	Sharpe		Sortino	
	(1)	(2)	(3)	(4)
	Key Urban Agglomerations	Non-Key Urban Agglomerations	Key Urban Agglomerations	Non-Key Urban Agglomerations
DFL	0.1365 *** (0.0063)	0.1595 *** (0.0087)	0.6802 *** (0.0665)	0.8787 *** (0.0963)
Control Variables	YES	YES	YES	YES
Year Fixed Effects	YES	YES	YES	YES
Province Fixed Effects	YES	YES	YES	YES
N	15377	10424	15377	10424
adj. R ²	0.2647	0.2339	0.0969	0.1054
Chow test P-value	0.0000		0.0000	

4.4.3 Heterogeneity Analysis by Household Financial Vulnerability

Given that household wealth conditions affect financial asset allocation, we partition the sample by financial vulnerability. We use the “household financial buffer” to capture financial vulnerability [26]. A negative buffer implies that a household cannot mobilize adequate funds to cope with financial hardship, thus falling into a financially vulnerable state; conversely, a non-negative buffer indicates resilience. The formula is specified as follows:

$$FM_{it} = HI_{it} + HLA_{it} - HC_{it} - HD_{it} \quad (11)$$

In Equation (11), FM_{it} represents the household financial buffer, defined as the contingency funds available to navigate financial hardship. HI_{it} is household total income. HLA_{it} is liquid assets held by the household that can be readily converted into cash, including time deposits, demand deposits, and cash. HC_{it} is total household consumption, which consists of expected and unexpected components. The former refers to routine living expenditures, whereas the latter primarily captures medical expenditures. HD_{it} is total household debt.

The subsample regression results in Table 10 indicate that digital financial literacy exerts a significant positive impact on the asset allocation efficiency of both types of households, but this enhancing effect is more pronounced among non-financially vulnerable households. A Chow test confirms that this cross-group difference is significant at the 1% level. The underlying reasons for this phenomenon may be twofold. On the one hand, non-financially vulnerable households, characterized by abundant assets and strong risk resilience, can leverage digital financial literacy to overcome information barriers and reduce transaction costs, thereby effectively utilizing digital tools for investment diversification and refined asset allocation. In contrast, vulnerable households are subject to liquidity constraints, with funds primarily prioritized to meet basic subsistence needs, making it difficult for digital financial literacy to be effectively translated into actual investment behavior. On the other hand, guided by risk control principles, financial institutions exhibit a preference for creditworthy and financially stable households. This enables the latter to more easily access low-cost financial resources, which further amplifies the asset optimization effect of digital financial literacy.

Table 10: Heterogeneity Analysis Results by Household Financial Vulnerability.

Variable	Sharpe		Sortino	
	(1)	(2)	(3)	(4)
	Non-Financially Vulnerable	Financially Vulnerable	Non-Financially Vulnerable	Financially Vulnerable
DFL	0.1444 *** (0.006)	0.0584 *** (0.0084)	0.753 *** (0.0632)	0.3111 *** (0.1054)
Control Variables	YES	YES	YES	YES

Year Fixed Effects	YES	YES	YES	YES
Province Fixed Effects	YES	YES	YES	YES
N	18325	7476	18325	7476
$adj.R^2$	0.2637	0.1804	0.1058	0.0693
Chow test P-value	0.0000		0.0000	

5. Conclusions and Recommendations

This study demonstrates that the positive impact of digital financial literacy on household financial asset allocation efficiency remains significant even after addressing endogeneity concerns and conducting robustness checks—including variable substitution and the exclusion of policy interference. Mechanism tests confirm that this effect operates primarily through two channels: improving household financial access and enhancing social interaction. Furthermore, heterogeneity analysis reveals significant group differences: the enhancing effect of digital financial literacy is more pronounced among rural households, those in non-key urban agglomerations, and non-financially vulnerable households, compared to their urban, key urban agglomeration, and financially vulnerable counterparts. Based on these findings, this paper proposes the following policy recommendations:

First, build a robust digital financial education system. Utilizing new media to popularize financial knowledge, the focus should be on improving residents' digital application and risk-awareness skills to foster rational investment. Second, optimize the digital financial service ecosystem. This involves advancing inclusive infrastructure coverage, incentivizing low-threshold and user-friendly products, and enhancing internet governance to eliminate misinformation, thereby converting financial literacy into actual asset allocation capacity. Third, precisely innovate inclusive products for the long-tail population. Moving beyond a “one-size-fits-all” model, institutions should be guided to develop agriculture-cycle-aligned products for rural areas, stratified service systems for lower-tier cities, and low-risk, high-liquidity emergency savings tools for vulnerable families.

References

- [1] Prasad H; Meghwal D; Dayama V. (2018). Digital financial literacy: a study of households of udaipur. *Journal of Business and Management*, 523–532.
- [2] Tony N; Desai K. (2020). Impact of digital financial literacy on digital financial inclusion. *International Journal of Scientific and Technology Research*, 9(1), 1911–1915.
- [3] Morgan P.J; Huang B; Trinh L.Q. (2019). The need to promote digital financial literacy for the digital age. In *Realizing Education for All in the Digital Age*, T20 Report, 40—46.
- [4] Wu W.X; Wu K; Zhang X.Y. (2018). Financial literacy and household portfolio effectiveness. *Journal of International Financial Studies*, (5), 66–75.
- [5] Wu W.X; Qiu Y.C; Zhang L.W. (2015). Effectiveness of Chinese household investment portfolios: based on the Sharpe ratio. *The World Economy*, 38(1), 154–172.
- [6] Pelizzon L; Weber G. (2009). Efficient portfolios when housing needs change over the life cycle. *Journal of Banking and Finance*, 33(11), 2110–2121.
- [7] Zang X.H; Dong J.X. (2023). Can social pension insurance improve the effectiveness of household portfolios?—a study from the life-cycle perspective. *Research on Economics and Management*, 44(10), 33–53.
- [8] Van R.M; Lusardi A; Alessie R. (2011). Financial literacy and stock market participation. *Journal of*

Financial Economics, 101(2), 449–472.

- [9] Wu Y; Li X; Li J, et al. (2021). Digital financial development and household financial portfolio effectiveness. *Management World*, 37(7), 92–104+7.
- [10] Wu W.X; Zhang X.Y; Wu K. (2021). Financial literacy and household saving rate: an explanation based on financial planning and borrowing constraints. *Journal of Financial Research*, (8), 119–137.
- [11] Zhou G.S; Liang Q. (2018). Internet use, market frictions, and household risky financial asset investment. *Journal of Financial Research*, (1), 84–101.
- [12] Wang X.H; Zhao Y.X. (2022). Digital financial development and the Urban-Rural disparity in household financial accessibility. *Chinese Rural Economy*, (1):44–60.
- [13] Liu S.J; Guan F.Q. (2024). Impact of digital financial literacy on financial asset allocation of rural households: from the perspective of financial availability. *Wuhan Finance*, (6), 57–64.
- [14] Wang X.H; Liu Y; Song M. (2023). Digital capability and household risky financial asset allocation. *Chinese Rural Economy*, (11), 102–121.
- [15] Si C.N; Zhang X.F. (2024). Digital financial literacy and household economic resilience—an analysis based on CHFS microdata. *Modern Finance and Economics (Journal of Tianjin University of Finance and Economics)*, (12), 52–67.
- [16] Jiang T; Li W; Sun W, et al. (2024). Digital financial inclusion and diversity of household financial asset allocation: evidence from China. *Australian Economic Papers*, 64(2), 205–220.
- [17] Chai S.J; Yang C; Lim Y.K. (2025). Digital divide and household portfolio investment efficiency: evidence from China. *Emerging Markets Review*, 70, 101407.
- [18] Calcagno R; Monticone C. (2015). Financial literacy and the demand for financial advice. *Journal of Banking and Finance*, 50, 363–380.
- [19] Li J; Alexis M. (2019). Promoting financial literacy through a digital platform: a pilot study in Luxembourg. *International Journal of Finance & Economics*, 26(1), 73–87.
- [20] Chai S.J. (2017). Social capital and household portfolio effectiveness. *China Economic Issues*, (4), 27–39.
- [21] Fu Q; Ma L.X; Hu J.L. (2022). Impact of social capital on financial asset allocation of rural households. *Economic Review*, (7), 118–128.
- [22] Liu C; Zhang Y.R; Li G.C. (2025). Digital financial capability and household wealth accumulation: evidence from the China Household Finance Survey. *Journal of Econometrics*, 5(2), 442–462.
- [23] Zhu W.P; Lin Y. (2022). Financial literacy, financial inclusion, and pension financial asset allocation. *Journal of Shanxi University of Finance and Economics*, 44(3), 43–57.
- [24] Ji, Y.L; Li, J.Y; Zhao H. (2022). Digital infrastructure construction and tourism economic growth: a mechanism test based on mediating and moderating effects. *On Economic Problems*, (7), 112–121.
- [25] Wang X.H; Li X.R; Song M, et al. (2024). Digital finance, digital divide, and household financial portfolio effectiveness: an analysis from the perspective of Urban-Rural disparity. *Contemporary Economic Science*, 46(2), 45–58.

- [26] Li B; Zhu T.H. (2022). Debt leverage, financial fragility, and heterogeneous household consumption behavior. *Journal of Financial Research*, (3), 20–40.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open - access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

The Impact of IP Holdings of Listed Gaming Companies on Investment Efficiency

Mada Su*

School of Economics and Management, Fuzhou University, Wulongjiang Avenue No. 2, Fuzhou 350108, China

**Corresponding author: Mada Su.*

Abstract

Recent years, the game industry has developed continuously. Game Intellectual Property (IP), as an important element in the game, has a very high popularity in the market, which has attracted more and more companies to concentrate on game IP for investment. However, there is a lack of multi-dimensional empirical research on the impact of game IP holdings on the investment efficiency of enterprises. It is found that the positive impact of IP holdings on inefficient investment has not passed the significance test, indicating that simply increasing the scale of IP holdings cannot significantly aggravate or alleviate the problem of inefficient investment; Enterprise growth has a significant inhibitory effect on inefficient investment, and high-growth enterprises are more likely to realize rational allocation of resources. This paper proves that it is difficult to convert the value of game IP into investment income alone, and blindly hoarding IP can easily lead to information asymmetry and aggressive investment by management, which in turn leads to resource mismatch. Learning the results of enterprise IP investment through this empirical analysis can improve the vigilance of enterprises against IP bubble, which can help to avoid blind investment in IP and turn their investment eyes to other sustainable growth points.

Keywords

investment efficiency, investment behavior of game companies, IP holdings, information asymmetry theory

1. Introduction

The full form of IP is “Intellectual Property”, which refers to highly focused, influential, and can be reproduced, encompassing creative intellectual property in creation [1]. In the game industry, IP, like traditional IP, can be extended based on the original characters and images, and its influence can be improved through artistic processing, collaborations, and so on.

Popular game IPs often attract the attention of many companies. Some popular IPs receive a large amount of capital injection continuously. Meanwhile, high premium valuations have also inflated the IP economy. More and more game companies are using well-established IPs to entice investors to put in large amounts of money, and investors are also willing to believe that developing products based on reliable IPs is an excellent way to reduce investment risk.

However, as the market’s enthusiasm for IP investment expands continuously, it has led some people to pay attention to the efficiency of IP investment. In recent years, some overexploited IPs have lost the favor of

players, causing what was once extremely popular to no longer be hot. For example, the ‘Need for Speed’ series, once the top global racing game IP, has gameplay with no innovation and uneven quality under the publisher’s ‘spring festival goods’ strategy, which made its reputation decline all the way. This study is based on analyzing the impact of IP holdings on the investment efficiency of gaming listed companies, specifically whether the continuous increase in IP holdings leads to efficient investment or inefficient investment, and whether higher IP holdings are more likely to result in overinvestment or underinvestment. Conducting such research can provide investors with a more scientific basis for decision-making, avoiding ineffective investments characterized by ‘high IP input, low content output,’ and breaking the industry misconception that ‘a famous IP is guaranteed to be a big hit’ [2]. It also corrects the industry’s cognitive bias of ‘IP-centrism,’ breaking the information bubble that ‘having an IP will automatically secure investment.’ This can effectively address the long-standing problem in the industry where ‘bad money drives out good money.’

2. Theoretical Basis and Research Hypotheses

From the perspective of information asymmetry theory, IP belongs to the typical category of intangible light assets, which are difficult to show in traditional financial statements in daily investments. Corporate investors and external investors often have information gaps, and investors generally only have fragmented information. Ha Thi Thu Nguyen and others pointed out that once speculative behaviors that exacerbate information asymmetry exist, they may lead to low investment efficiency [3]. While gaming companies acquire a large amount of IP through investment and acquisitions, because IP as an intangible asset is hard to quantify, it is more likely to exacerbate information asymmetry and ultimately lead to a decline in the company’s investment efficiency.

The theory of intangible assets further supplements the cyclical characteristics of IP. Arrighetti, Alessandro also indicated that intangible assets go through an accumulation process, during which some heterogeneity phenomena may occur, with both underinvestment and overinvestment being possible [4]. Due to the longer investment cycle of intangible assets, listed gaming companies are more likely to exhibit heterogeneity in their pursuit of popular IPs, which in the short term manifests as underinvestment by the companies. Based on the principal-agent theory, the traffic bonus of popular IPs is significant. The management may be prompted to adopt more aggressive investment behaviors due to considerations such as industry reputation, and ignoring the compatibility between the IP and the company itself may lead to resource misallocation, ultimately leading to a decline in investment efficiency. Garcia, R., & Silva, M. empirically demonstrated that the scale of holding intangible assets is significantly negatively correlated with a company’s investment efficiency. The higher the proportion of intangible assets and the larger the total holdings, the more severe the company’s inefficient investments. Due to its characteristics that are difficult to quantify, it exacerbates information asymmetry and principal-agent conflicts, ultimately resulting in underinvestment or overinvestment. Their results on industry heterogeneity also clearly indicate that compared to traditional physical industries, IP has a more pronounced positive effect on inefficient investment [5]. Anaka, S., & Sato, K analyzed the phenomenon of overinvestment caused by management blindly following the trend of popular IP due to information asymmetry for purposes such as performance evaluation, thereby neglecting the company’s own R&D capabilities and leading to improper resource allocation [6]. Companies with high IP holdings generally exhibit phenomena of underinvestment or overinvestment. Based on the above theory, this paper proposes the following assumptions.

H1: The amount of IP holdings has a significant impact on the inefficient investment of listed gaming companies.

3. Research Design

This study chose 20 gaming companies as samples from A-shares during 2020-2025, excluding ST stocks, financial companies, and samples with missing key data. The samples come from the relevant financial report information in the Guotai An database. Missing data are discarded according to the case-by-case exclusion method.

3.1 Variable Definition

The explanatory variables in this paper are IP holdings (IP) and inefficient investment (UNIV). According to Liu Tian and Xiang Xianhu, under the constraints of current accounting standards, intangible assets other

than intellectual property are excluded from the financial accounting system, manifesting as off-balance-sheet intangible assets [7]. According to the provisions of the new accounting standards, only intellectual property intangible assets can be converted into intangible assets. Therefore, this paper uses the proportion of intangible assets to total assets to measure the amount of IP held. Inefficient investment is the absolute value of the non-zero residuals obtained from the Richardson model. The dependent variables in this article are overinvestment (OverInv), corporate growth (Growth), debt ratio (Lev), cash holdings (Cash), and the proportion of independent directors (Ind), and overinvestment in the Richardson model is defined as a residual greater than zero. The control variables in this article are executive compensation (Pay), administrative expense ratio (Adm), management shareholding ratio (M__SHDS), and major shareholder appropriation ratio (Orecta).

3.2 Model Setup

$$UNIV = \alpha_0 + \alpha_1 IP + \alpha_2 Growth + \alpha_3 Lev + \alpha_4 Cash + \alpha_5 Ind + \alpha_6 Pay + \alpha_7 Adm + \alpha_8 M\_SHDS + \alpha_9 Orecta + \varepsilon \quad (1)$$

4. Empirical Results Analysis

4.1 Descriptive Statistics

Table 1: Variable Distribution

Variable	(Mean)	(Std. dev.)	(Min)	(Max)
UNIV	0.1296374	0.1397566	0.0000664	0.802322
IP	0.031552	0.0388462	0.0000902	0.181546
Growth	2403.753	19078.76	-53.13565	151433
Lev	0.2883843	0.13453	0.0735399	0.6408477
Cash	0.2044957	0.1504963	0.0033083	0.6550932
Ind	41.23353	4.175627	33.33	50
Pay	15.51548	0.801928	14.07015	17.58481
Adm	2.715399	16.17122	0.014307	139.6461
M SHDS	16.63286	14.63557	0	45.6544
Orecta	0.0225509	0.0376927	0.0004314	0.2210763

As shown in Table 1, inefficient investment (UNIV) is measured by the absolute value of the residuals from the Richardson model, with a mean of 0.1296, a standard deviation of 0.1398, a minimum value of 0.0001, and a maximum value of 0.8023, indicating that the overall level of inefficient investment among the sample firms is relatively low, but there are certain differences in investment efficiency among the firms. The effective average of IP holdings (IP) is 0.0316, with a standard deviation of 0.0388, reflecting that the IP holdings of sample listed gaming companies are generally low and relatively concentrated. The average corporate growth (Growth) is 2403.753, with a standard deviation as high as 19078.76, showing a huge disparity in extreme values. This indicates that the growth of sample enterprises is significantly differentiated, with some companies demonstrating outstanding growth potential while others face growth bottlenecks. The average asset-liability ratio (Lev) is 0.2884, with a standard deviation of 0.1345 and a range of 0.0735 to 0.6408, indicating that the financial leverage of the sample companies is generally conservative and the debt repayment risk is controllable.

The average cash holding level (Cash) is 0.2045, which aligns with the industry characteristic of higher cash reserves among A-share gaming companies. The administrative expense ratio (Adm) has extreme values (maximum 139.6461), and the minimum management shareholding ratio (M_SHDS) is 0, indicating that some companies' management does not hold shares, and there are differences in governance structures. The proportion of independent directors (Ind), executive compensation (Pay), and the proportion of other receivables (Orecta) all exhibit distribution patterns consistent with the characteristics of the gaming industry. In addition, some variables (such as UNIV and Growth) have missing observations, which is related to the sample selection of the Richardson model during the data matching process. Subsequent regression analysis uses balanced panel data to ensure the reliability of the results.

To test the impact of IP holdings on inefficient investment (underinvestment) of listed gaming companies, this paper sequentially constructs an OLS benchmark model (Model 1) and a fixed-effects panel model (Model 2), with the regression results shown in Tables 2 and 3.

Table 2: Baseline Regression

Variable	(1) OLS UNIV	(2) FE UNIV
IP	0.507(0.368)	3.921(2.718)
Growth	-0.00000379(0.00000245)	-0.00000341**(0.00000127)
Lev	-0.292(0.212)	-1.032(0.857)
Cash	0.154(0.242)	-0.0421(1.529)
Ind	0.00170(0.00588)	0.0112(0.00896)
Pay	-0.00396(0.0189)	-0.0294(0.128)
Adm	0.000172(0.000486)	0.0000992(0.000458)
M_SHDS	0.000687(0.00142)	-0.000977(0.00290)
Orecta	0.826(0.629)	2.796(4.895)
cons	0.120(0.334)	0.241(1.985)
N	30	30

Table 3: Regression Analysis

Variable	Coefficient	Robust std. err.	Tvalue	P value
IP	3.921027	2.717624	1.44	0.18
Growth	-3.41E-07	1.27e-07	-2.68	0.023**
Lev	-1.032389	0.8573306	-1.2	0.256
Cash	-0.0420594	1.528777	-0.03	0.979
Ind	0.0112473	0.0089581	1.26	0.238
Pay	-0.0294034	0.1276255	-0.23	0.822
Adm	0.0000992	0.0004576	0.22	0.833
M_SHDS	-0.0009768	0.002902	-0.34	0.743
Orecta	2.796194	4.895118	0.57	0.58
cons	0.2407542	1.984812	0.12	0.906
sigma u	0.17207031			
sigma e	0.07908151			

In Table 2, Model (1) presents the OLS benchmark regression results. The coefficient of the core explanatory variable, IP, is 0.507, with a t-value of approximately 1.38 and a p-value greater than 0.1, failing to pass the 10% significance level test. This indicates that, without controlling for individual heterogeneity, IP holdings have no significant effect on inefficient investment, which is inconsistent with the expectation of Hypothesis 1 in this paper. Among the control variables, firm growth has a t-value of about -1.55 and is not significant; the remaining control variables (Lev, Cash, Ind, etc.) are also not significant. The overall explanatory power of the model is limited, reflecting that the model specification is biased when individual effects are not controlled.

To control for firm-level individual heterogeneity (such as unobservable factors like corporate culture and governance models), this paper uses a fixed effects model (FE) for regression.

The coefficient of IP holdings (IP) is 3.921 (standard error 2.718), with a t-value of 1.44 and a P-value of 0.180, still failing to pass the 10% significance level test. This indicates that after controlling for individual fixed effects, IP holdings still have no significant impact on inefficient investment. Therefore, Hypothesis 1 of this paper, "IP holdings have a significant negative impact on underinvestment in listed gaming companies," is not yet supported by the data.

The coefficient of corporate growth (Growth) is -3.41e-07 (standard error 1.27e-07), with a t-value of -2.68 and a p-value of 0.023, which is significantly negative at the 5% level. This indicates that the higher the corporate growth, the lower the level of inefficient investment, which is consistent with the results of the correlation analysis, confirming the inhibitory effect of growth on corporate investment efficiency, that is, high-growth companies are more likely to allocate resources reasonably and reduce underinvestment.

The P-values of asset-liability ratio (Lev), cash holdings (Cash), proportion of independent directors (Ind), executive compensation (Pay), administrative expense ratio (Adm), management shareholding ratio (M_SHDS), and proportion of other receivables (Orecta) are all greater than 0.1, indicating no stable significant effect on inefficient investment.

In fact, because game IPs are intangible assets, their value is highly uncertain. Therefore, a high holding of IP by a company cannot guarantee efficient investment, resulting in data that cannot pass the 10% significance

test. It may also be because the game industry, compared to traditional physical industries, has a longer development cycle, making heterogeneity more likely, which leads to inefficient investment. Moreover, the experimental time span is not long, and in a short period, the holding of IP cannot quickly translate into effective investment returns, making the hypothesis invalid.

5. Suggestions

Based on the research results of this article, the following suggestions are proposed. Game companies should view the value of IP rationally and should not use it as the core basis for investment decisions. Companies should weaken the investment decision logic of ‘only IP matters.’ Game companies that pursue the quantity of popular IPs too much while neglecting content development are more likely to make inefficient investments. Gaming companies should place more emphasis on establishing an investment rating system centered on research and development capabilities, product quality, and operational efficiency. This is supported by Oguguo’s findings: Oguguo conducted an empirical analysis of the operational data and investment behaviors of 200 gaming companies of different sizes, and the results indicated that IP assets themselves do not have the ability to be converted into operational benefits and investment returns; the realization of IP value relies more on the company’s own research, development, and innovation capabilities [8]. This conclusion directly refutes the ‘IP-only theory’ and demonstrates that the realization of IP value requires enterprises to deeply integrate IP value with innovation capabilities and technology application. Enterprises can also avoid inefficient investments caused by blindly pursuing popular IP by comprehensively evaluating multiple aspects such as the IP lifecycle, content innovativeness, and market acceptance. Meanwhile, companies can also reduce agency costs and inefficient investments by strengthening their corporate governance. The research by Wang Y. and Li S. confirmed the inhibitory effect of corporate governance mechanisms on inefficient investments. It indicates that a reasonable equity balance structure effectively restrains the behavior of major shareholders in occupying funds, transferring benefits, and blindly leading investments, and can also reduce the agency conflicts between shareholders and minority investors. Kweh, Q. L.’s research verified the inhibitory effect of corporate governance mechanisms on inefficient investments. It showed that a reasonable equity balance structure effectively restrains major shareholders from occupying funds, transferring benefits, and blindly leading investments, and can also reduce agency conflicts between shareholders and minority investors. Kweh, Q. L. also pointed out that strengthening incentive and constraint mechanisms for management is crucial for improving corporate investment efficiency [9]. Therefore, companies should further optimize the board structure, strengthen the board’s supervisory functions, and mitigate agency problems through reasonable management shareholding and compensation contracts. They can also try to establish a binding incentive mechanism that links managerial compensation stock closely with the company’s long-term performance. At the same time, strengthen the supervision of major shareholders’ use of funds to prevent the chaos of irrational investment expansion, reduce insider control and resource misallocation, thereby improving the overall investment efficiency and capital allocation effectiveness of enterprises.

Enterprises can also enhance investment efficiency through strengthening growth management by driving it with business growth. The core conclusion of Hsiung et al.’s research shows that the business growth of games is significantly positively correlated with investment efficiency. By focusing on extending core categories, promoting the layout in emerging overseas markets, or deepening the development of the entire IP industry chain, enterprises improve overall utilization efficiency, avoid the flow of funds to low-return, inefficient projects, and thereby directly suppress inefficient investment. Additionally, research has also shown that once the growth level crosses a critical threshold, the optimization effect of corporate capital allocation will be further amplified [10]. Therefore, game companies should focus on enhancing innovation and market expansion in their core businesses. Since high growth can significantly suppress investment efficiency, companies should allocate resources to projects with long-term growth potential, driving continuous improvement in investment efficiency through high-quality growth.

6. Conclusion

This article studies the relationship between IP holdings and inefficient investment caused by the low investment efficiency of gaming companies listed due to the ‘IP-only theory’ phenomenon. This article uses

data from Chinese listed gaming companies and finds that IP holdings have a significant effect on inefficient investment. It proves that a company's investment model should not focus solely on improving investment efficiency through a high IP ownership, but can also alleviate principal-agent problems by improving internal corporate governance, optimizing board structure, strengthening management incentives and constraints, and establishing reasonable management shareholding and compensation contracts. In the same time, IP should be combined with the company's innovation capabilities. However, this paper still has limitations. The sample size of companies selected in this paper is small, which may result in the results being accidental. In addition, this paper did not test for multicollinearity, which may lead to deviations in the judgment of the impact of each factor on investment efficiency, further weakening the rigor of the experiment. In the future, the sample design will be comprehensively optimized to overcome the shortcomings of small-sample studies, while extending the research period to enhance the rigor of the conclusions. Future research can further explore in-depth aspects. By introducing mediating variables such as management decisions and technological innovation, the specific pathways through which IP affects inefficient investment can be investigated, clarifying the logical relationship between the two. Additionally, different types of IP, such as original IP and licensed IP, can be discussed separately to make the research findings more precise.

References

- [1] Wang Zhong, Song Shaoting. Analysis of the Correlation between Investment Efficiency of Listed Companies in the Cultural and Media Industry [J]. *China Culture Industry Review*, 2020, 28(1): 71-96.
- [2] Yan Xinyue, Deng Xiangyang. Breaking the Myth of IP: Analysis of the Impact of IP on Film and Television Investment Efficiency [J]. *Media Economics and Management Research*, 2022, (03): 191-210.
- [3] Nguyen H T T, Nguyen H T T, Pham A T, et al. Earnings Management, Investment Efficiency, and High Information Uncertainty: Evidence from Vietnam[C]//International Symposium on Integrated Uncertainty in Knowledge Modelling and Decision Making. Singapore: Springer Nature Singapore, 2025: 195-208.
- [4] Arrighetti A, Landini F, Lasagni A. Intangible assets and firm heterogeneity: Evidence from Italy[J]. *Research Policy*, 2014, 43(1): 202-213.
- [5] Ahmed M S, King T. The dark side of intangibles? Organizational capital and corporate investment efficiency[J]. *Journal of Accounting Literature*, 2025, 47(5): 444-489.
- [6] Hayashishita E, Oda T. Industry differences in intellectual property disclosure and corporate value: Evidence from Japan. Access to science, business, innovation in digital economy[J]. ACCESS Press, 2025, 6(3): 498-512.
- [7] Xiang Xianhu, Liu Tian. On off-balance-sheet intangible assets: From the perspective of integrating finance and strategy—also analyzing intangible resources, intangible assets, and intangible capital [J]. *Accounting Research*, 2014, (04): 3-9 95.
- [8] Oguguo P C. Innovation and intellectual property use in the global video game industry[J]. *World Intellectual Property Organization (WIPO) Economic Research Working Paper Series*, 2024 (85).
- [9] Kweh, Q. L., Ting, I. W. K., Lu, W. M., & Le, H. T. M. (2025). Controlling Interest and Firm Efficiency: Comparing Family-and Nonfamily-Controlled Taiwanese Cultural and Creative Industry Firms. *Managerial and Decision Economics*, 46(2), 1344-1357.
- [10] Hsiung H H, Lin C Y, Zhu G Y. The impact of intellectual capital efficiency on value creation in the video game industry—Evidence from Taiwan[J]. *Journal of Infrastructure, Policy and Development*, 2023, 7(3): 1-17.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

The Alienation of Behavioral Biases Among A-Share Investors from a Localization Perspective

Zekai Jiang*

School of Management Science and Engineering, Central University of Finance and Economics, Beijing 100000, China

**Corresponding author: Zekai Jiang.*

Abstract

Behavioral finance has established a complete theoretical system in mature Western capital markets, effectively explaining the alienation induced by Market Behavioral Biases. However, the A-share as an emerging policy-driven and retail-driven market, has a unique investor structure, trading mechanism, and indigenous sociopsychological characteristics, leading to a significant alienation of classic behavioral bias theory. This article will take the Alienation of Behavioral Biases Among A-Share Investors as the core object of study. And based on Prospect Theory and employing a literature review approach, this article will conduct a systematic analysis of such aberrant phenomena as the polarization and reversal of the Disposition Effect, the full-chain extension of the Herd Effect, and the cyclical reversal of overconfidence, also delve deeply into the roots of distortion—including trading systems, information asymmetry, the retailization of institutional investors, and local culture. Research indicates that Western behavioral finance theory is not fully adapted to the unique institutional and cultural environment of the A-share market, unable to explain its underlying causes. Promoting the localization of behavioral finance theory is a critical path for correcting A-share investor behavioral alienation and enhancing market pricing efficiency, providing a theoretical basis for regulatory policy formulation and investor strategy construction. This paper addresses the gap in traditional research that overlooks institutional and cultural differences in emerging markets, while acknowledging limitations such as the lack of empirical testing. Future research could further refine the localized behavioral finance theoretical system through empirical and quantitative studies.

Keywords

localization, A-share market, investor behavioral biases, behavioral alienation, behavioral finance

1. Introduction

In recent years, behavioral finance has established a comprehensive theoretical framework in mature Western capital markets, effectively explaining anomalies stemming from behavioral biases in market conduct [1]. However, the A-share market—characterized as an emerging market driven by both policy and retail investors—exhibits a unique investor structure, distinct trading mechanisms, and specific indigenous sociopsychological traits; consequently, classic Western theories of behavioral bias undergo significant divergence when applied within this context. This alienation is not only reflected in macro-level phenomena, including the extremization and reversal of the disposition effect [2, 3], the full-chain propagation of the herd effect [4, 5], and the cyclical reversal of overconfidence [6, 7]; but also further profoundly rooted in micro-level

mechanisms and socio-psychological factors, including the T+1 trading system, price limit mechanisms, information asymmetry, the “retailization” of institutional investors, and indigenous Chinese culture [8]. In light of this, this paper aims to conduct an in-depth analysis of the phenomenon of behavioral biases among A-share investors from a localized perspective. It seeks to reveal the limitations of Western behavioral finance theories when applied to the specific market environment of China, while emphasizing both the necessity of—and the pathways toward—adapting behavioral finance theory to the local context. This study not only makes up for the deficiencies of traditional behavioral finance research—which overlooks the institutional disparities and indigenous cultural characteristics of emerging markets—but also refines the theoretical framework of behavioral finance with Chinese characteristics. It also provides theoretical foundations and practical guidance for regulatory bodies to formulate policies aligned with market realities and for institutional and individual investors to build strategies adapted to the local market. Furthermore, it offers a reference paradigm for the localization of behavioral finance research and market governance in other emerging capital markets.

2. Background

The fundamental objective behind the establishment of China's A-share market was to facilitate the reform of state-owned enterprises and to construct a platform for financing and investment that connects corporations with investors. As a market that is both policy-driven and retail-dominated, the A-share market is characterized by the psychological traits of its investor base and a massive number of retail investors [9]. With its continued evolution, the A-share market currently displays a landscape dominated by retail investors, marked by the rise of private equity funds, while the proportions held by general corporate entities and state-backed “national teams” remain relatively stable [10]. Within this structure, the massive population of retail investors exerts significant influence yet remains unable to effectively contend with institutional investors. The core issue underlying this is information asymmetry, primarily reflected in information disclosure, intermediation, dissemination, supervision, and the information-processing capabilities of investors [11]. Currently, the trading mechanisms of China's A-share market remain immature; the T+1 and price limit mechanisms constrain price adjustments to a certain extent [8], exacerbating volatility and lag, and making individual investors more susceptible to external influences. It is precisely due to these distinct localized characteristics of the A-share market that existing behavioral bias theories have undergone alienation within this market—a form of alienation that differs from other capital markets and necessitates more localized research for explanation.

3. Research Objective

Given the unique institutional and cultural context of the A-share market, investor behavioral biases exhibit idiosyncratic characteristics that diverge significantly from Western theoretical frameworks. To gain a comprehensive understanding of—and effectively address—these idiosyncratic phenomena, this study is structured around the following three core objectives:

Research Objective 1: To grasp the overall landscape of behavioral biases within the A-share market.

Research Objective 2: Building upon these macro-level characteristics, analyze the micro-level psychological and decision-making mechanisms behind behavioral biases.

Research Objective 3: Identify effective pathways for the connection between macro manifestations and micro mechanisms.

For the first research objective, it is essential to examine their idiosyncratic manifestations from a macro perspective, thereby elucidating the macro-level idiosyncratic features of behavioral biases in this market. Regarding the second objective, the study will delve into the micro perspective by dissecting the psychological cognition, decision-making processes, and market interaction patterns of individual investors, thereby providing an in-depth analysis of the micro-level behavioral factors underlying these idiosyncratic phenomena. The third objective represents the ultimate purpose of understanding the connection between macro manifestations and micro mechanisms. This study will explore the necessity of correcting idiosyncratic behaviors within the A-share market and analyze the potential theoretical and practical implications of such interventions, with the aim of offering policy recommendations to foster the healthy development of the market.

4. Theoretical Framework

Prospect theory, proposed by Kahneman and Tversky [1], holds that individuals' decision-making under uncertainty is not based on absolute utility, but on changes in gains and losses relative to a reference point, and exhibits significant asymmetric characteristics. The core principle would present as follow. First, loss aversion, where the degree of pain caused by an equivalent loss is significantly greater than the pleasure derived from an equivalent gain, leading individuals to have extremely high negative sensitivity to risk perception. Second, risk preference reversal is manifested as a tendency toward risk aversion (settling for sure gains) when facing profits, and a tendency toward risk seeking (taking a “last-ditch” gamble) when facing losses. Finally, diminishing sensitivity, meaning that as the absolute value of gains or losses increases, the marginal psychological utility they produce gradually weakens.

5. Literature Review

5.1 Market-driven Upheavals in Prospect Theory—The Extremes of the Disposition Effect

In classical prospect theory, investors tend to overestimate low-probability events. However, under the A-share price limit system, this psychological distortion is forcibly amplified by external institutions. When the stock price is close to the price limit, the market will exhibit a significant “magnetic effect”: the price limit will induce investors' “fear of missing out”, causing them to subjectively overestimate the probability of the stock price continuing to rise, thus leading to irrational chasing behavior near the price limit [12]; conversely, the price limit will cut off liquidity, causing investors to have an extreme panic of “liquidity depletion”, thus overestimating the probability of subsequent crashes. This institutionalized “probability induction” directly alters the probability weighting function curve in Prospect Theory, making it exhibit steeper non-linear characteristics at the boundaries compared to Western markets.

The T+1 trading system is a liquidity constraint unique to the A-share market, which directly interferes with the settlement cycle of investors' “mental accounts”. Under the T+1 constraint, investors who buy on the same day cannot perform “corrective” selling. This makes the investor's reference point no longer just the purchase cost, but a dynamic anchor point superimposed with “time premium” and “immediate liquidity risk” [13]. Studies have found that this mandatory holding and waiting strengthens the investor's “loss aversion” coefficient: due to the inability to stop losses in time, the psychological pressure on investors increases exponentially when facing intraday fluctuations. Loss aversion in this institutional environment is not a static personality trait but a form of “institutional panic” constrained by trading rules, leading investors to be extremely conservative when in profit, while falling into “learned helplessness” when in loss due to the inability to take immediate action, ultimately evolving into a deep alienation of the disposition effect.

The most significant dramatic change is the “disposition effect reversal” phenomenon that occurs in extreme game scenarios in A-shares (such as new stock speculation and consecutive limit-up stocks). Traditional theory predicts that investors will “sell winners and hold losers,” but in the high-frequency gaming of A-shares—characterized by high retailization and severe information asymmetry—investors generally exhibit a “lottery-like preference” [14]. Under the protection of the limit-up system, profitable positions are regarded as a scarce “entry ticket”. Investors not only do not take profits, but also have a strong “gambler's fallacy” and continue to add positions. While when facing consecutive limit-downs, due to the liquidity trap under the T+1 system, investors often choose to “panic sell” rather than continue to hold when they can sell [15]. This reverse behavior of “selling losers and holding winners” is essentially a parameter alienation of Prospect Theory under the combined effect of extreme liquidity constraints and local gaming psychology, reflecting a regression of A-share investors' decision-making logic from “value judgment” to “emotional survival” under institutional pressure.

5.2 The Cyclical Escalation of the Herding Effect and the Cyclical Reversal of Overconfidence

The disposition effect has further spurred a more prominent full-chain herding effect in the A-share market. This is an inevitable upgrade of individual behavioral deviations to group behavioral deviations through market interaction. The classic herd effect theory distinguishes between rational herding based on information asymmetry and irrational herding based on emotional imitation [4]. However, A-share herding behavior is

highly decoupled from corporate fundamentals: while mature market investors use intrinsic corporate value as a decision reference point, A-share investors' reference points have systematically alienated into “others' holding returns” and “theme popularity” [16, 17]. At the same time, the positive feedback mechanism of “chasing highs and selling lows” brought about by the disposition effect has led to the rapid spread of the profit-making effect of theme speculation. An “institutions first - large investors follow - retail investors follow the trend” imitation trading model has emerged in the whole chain, making investors' holdings highly similar and the track extremely clustered, which has been distorted into a pure game of emotional speculation on the rise and fall limit [5, 18].

The periodic outbreak of the herding effect in the group, in turn, distorts the self-cognition of investors, forming the local alienation characteristic of overconfidence. This is the reverse shaping process of group behavior on individual psychology. Classic overconfidence theory considers investor overconfidence as a stable personality trait and cognitive bias [6]. However, A-share investors' overconfidence exhibits extreme periodic reversal characteristics. In the bull market cycle, the general short-term profits brought by the herding effect will be mistakenly attributed by investors to their own game ability. At the same time, the continuously rising stock price will continuously raise the decision reference point of investors, making them overestimate their own ability to control risks, and eventually evolve into the extreme radicalism of the whole people adding leverage to enter the market [7]. In the bear market cycle, the general losses brought by the herding effect will be simply attributed by investors to the external environment. At the same time, the rapidly falling stock price will greatly reduce their decision reference point, making them instantly underestimate their own investment ability. Overconfidence directly reverses into the extreme conservatism of panic selling [19, 20].

5.3 Reasons for Mechanism Alienation from an Economic Perspective

The aforementioned alienation of behavioral biases in the A-share market is essentially a parameter reconstruction and mechanism mutation of classic behavioral finance models under the specific micro-environment of A-shares. Its roots span the entire investor decision-making chain and are superimposed with multiple local factors: At the information processing stage, investors exhibit significant attention-salience bias and lottery-like information screening preferences. The echo-chamber effect of social media and viral information dissemination further solidifies the micro-foundation of the herd effect [14]. At the cognitive judgment stage, the T+1 and price limit systems distort the risk preference parameters of Prospect Theory, leading to intensified loss aversion asymmetry and the extremization of the disposition effect. Simultaneously, overconfidence and historical price anchoring effects are prevalent, with investors generally relying on intuitive systems for decision-making while suppressing rational analysis [21, 22]. At the trading execution stage, the T+1 system reinforces regret aversion and the sunk cost fallacy, while the magnet effect of price limits fosters a pattern of consecutive-limit speculation [23]. Moreover, the extremely weak short-selling mechanism creates a one-sided market structure, making it difficult to correct overvalued mispricing [24]. Institutional investors, who should serve as market stabilizers, exhibit a “retailization” tendency under the pressure of short-term performance rankings and principal-agent problems; their behavior of “clustering for warmth” (herding) and sector rotation instead amplifies market volatility [5], while trend-following strategies in quantitative trading further exacerbate the effects of “helping the rise and adding to the fall.” Additionally, China's local tradition of collectivism and the generally low financial literacy of investors provide deep socio-psychological soil for these behavioral biases, ultimately causing A-shares to exhibit systematic irrational characteristics that are more extreme in degree, longer in duration, and more unique in form.

6. Key Finding

This study delves into the alienation of investor behavior biases in the A-share market from a localized perspective, revealing the limitations of Western behavioral finance theory in explaining the specific market environment of China. The research finds that A-share market behavioral biases exhibit significant macro-level alienation characteristics, specifically manifested in the extremization and reversal of the disposition effect, the full-chain nature of the herd effect, and the periodic reversal of overconfidence. These alienated behaviors are rooted in the unique micro-mechanisms and local factors of the A-share market, including the distorting effect of the T+1 trading system and price limit mechanism on the risk preference parameters of prospect theory, information asymmetry and cognitive biases, as well as the social and psychological factors such as the tendency of institutional investors to become retail investors and China's collectivist tradition.

Therefore, this study emphasizes that promoting the localization and “Chinese-characterization” of behavioral finance theory is a key path for correcting the alienation of A-share investor behavior and enhancing market pricing efficiency. This not only provides a theoretical basis and practical guidance for regulatory departments and investors but also fills the gap in traditional behavioral finance research and provides a reference paradigm for localized behavioral finance research in other emerging capital markets.

7. Discussion

7.1 Applications and Significance

Localizing and adapting behavioral finance theory to Chinese characteristics to correct market anomalies has both clear practical application value and significant academic and market implications. In terms of application, this research approach directly provides adapted analytical tools for interpreting unique behavioral biases such as the reversal of the disposition effect and the full-chain herd effect in A-shares. It accurately depicts the investor decision-making logic under the combined influence of local socio-psychology and special trading rules. This can provide a theoretical basis for regulatory authorities to formulate behavior-guiding and rule-optimizing policies that align with market realities, and offer practical guidance for institutional and individual investors to build investment strategies adapted to the local market and avoid irrational alienated behaviors. In terms of significance, this approach fills the gap in traditional behavioral finance research by neglecting the institutional differences and local cultural characteristics of emerging markets. It improves the theoretical system of behavioral finance with Chinese characteristics, breaks through the limitations of Western theories in explaining real-world market anomalies in the A-share market, and lays a solid theoretical foundation for systematically correcting A-share investor behavior distortions, improving market pricing efficiency, and maintaining market stability. This further contributes to the deepening of the registration-based reform and high-quality development of the capital market, and also provides a reference paradigm for the localization of behavioral finance research and market governance in other emerging markets.

7.2 Research Shortcomings and Research Directions

The paper also has some shortcomings: it relies solely on literature review for theoretical analysis, without conducting empirical tests, quantitative modeling, or case studies, resulting in conclusions lacking data support. Future research could focus on conducting more empirical tests and quantitative analyses of the formation and correction mechanisms of behavioral biases, continuously improving and applying localized behavioral finance theory, and laying a solid theoretical foundation for the deepening of the registration-based reform and high-quality development of China’s capital market.

8. Conclusion

From a localized perspective, this paper takes the alienation of A-share investor behavioral biases as its core research object. Based on Prospect Theory and Limited Arbitrage Theory [25], and combined with the unique trading systems, investor structure, and local socio-psychological characteristics of the A-share market, it systematically reviews the alienated manifestations, micro-formation mechanisms, and deep-seated causes of typical behavioral biases such as the disposition effect, herd effect, and overconfidence. The paper further finds that behavioral finance theories from mature Western markets are difficult to adapt to the emerging policy-driven and retail-dominated A-share market, as they do not fully incorporate local special trading rules and socio-cultural factors, and thus cannot effectively explain the unique behavioral bias alienation phenomena in A-shares. Promoting the localization and “Chinese-characterization” of behavioral finance theory is the core path for correcting the alienation of A-share investor behavior. This transformation can provide adapted analytical tools for interpreting the decision-making logic of local investors, offer theoretical support for regulators to optimize market mechanisms and guide rational investment, and assist institutional and individual investors in avoiding irrational behaviors and formulating investment strategies adapted to the local market. This study fills the gap in traditional behavioral finance research that neglects institutional differences and local cultural characteristics in emerging markets, refines the behavioral finance theoretical system with Chinese characteristics, and provides a reference paradigm for localized behavioral finance research and market governance in other emerging capital markets globally.

References

- [1] Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291. <https://doi.org/10.2307/1914185>
- [2] He, J., & Wang, L. (2023). Evolution of A-share investors' disposition effect and institutional influence under registration system reform. *Securities Market Herald*, 33(9), 4–13. <https://doi.org/10.19622/j.cnki.cn31-1506/f.2023.09.001>
- [3] Zhang, Q., Xie, C., & Cheng, T. (2024). The impact of daily-frequency investor sentiment on A-share short-term returns. *Operations Research and Fuzziology*, 14(2), 163–179.
- [4] Bikhchandani, S., & Sharma, S. (2001). Herd behavior in financial markets. *IMF Staff Papers*, 47(3), 279–310.
- [5] Xu, N., Yu, S., & Yi, Z. (2013). Institutional investors' herding behavior and stock price crash risk. *Management World*, 29(7), 31–43. <https://doi.org/10.19744/j.cnki.11-1235/f.2013.07.004>
- [6] Barber, B. M., & Odean, T. (2001). Boys will be boys: Gender, overconfidence, and common stock investment. *The Quarterly Journal of Economics*, 116(1), 261–292. <https://doi.org/10.1162/003355301556099>
- [7] Yang, C., Wu, C., & Chen, M. (2005). Investors' non-persistent overconfidence and stock market reversal effect. *Management Science*, 18(2), 40–45.
- [8] Xiao, X. (2024). China's A-share price limit trading mechanism and investors' disposition effect. *Journal of Financial Research*, (9), 123–140.
- [9] Leng, S. (2016). Analysis of policy-oriented characteristics of A-share investment and countermeasures. *Journal of Yangzhou Polytechnic Institute*, 20(3), 23–26. <https://doi.org/10.15954/j.cnki.cn32-1529/g4.2016.03.006>
- [10] Wang, R. (2018). Major changes in the investor structure of China's capital market. *Financial Economy*, (8), 99–101. <https://doi.org/10.14057/j.cnki.cn43-1156/f.2018.08.042>
- [11] Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *American Economic Review*, 70(3), 393–408.
- [12] Fudan University International Finance Institute. (2026, March 20). The “magnet effect” of the price limit system and market microstructure distortion. <https://fisf.fudan.edu.cn/ffr/content/314> (Accessed April 12, 2026)
- [13] Guo, M., Liu, T., & Wu, M. (2019). Should the “T+1” trading system be abolished? *Tsinghua Financial Review*, (10), 81–83. <https://doi.org/10.19409/j.cnki.thf-review.2019.10.019>
- [14] Wang, Y., He, J., & Zang, R. (2018). The impact of short-sale constraints on securities market mispricing. *Journal of Beijing University of Aeronautics and Astronautics (Social Sciences Edition)*, 31(1), 81–88.
- [15] Zhang, X. (2023). The generation mechanism of major risks in China's stock market—From the perspective of liquidity supply-demand imbalance. *Economic Management Journal*, (2), 45–58.
- [16] Song, J., & Wu, C. (2001). A study of herding behavior in financial markets based on dispersion. *Economic Research Journal*, 46(11), 21–27+93. <https://doi.org/10.19744/j.cnki.11-1081/f.2001.11.003>
- [17] Shi, D. (2001). Trading behavior of securities investment funds and its market impact. *World Economy*, 24(10), 26–31. <https://doi.org/10.19985/j.cnki.cassjwe.2001.10.004>
- [18] Qi, B., Huang, M., & Chen, Z. (2006). Institutional investors and stock market volatility. *Journal of Financial Research*, 32(9), 54–64. <https://doi.org/10.19795/j.cnki.cn11-1262/f.2006.09.006>
- [19] Zhao, X., & Wang, Y. (2001). An empirical analysis of the “disposition effect” in China's stock market. *Journal of Financial Research*, 27(7), 92–97. <https://doi.org/10.19795/j.cnki.cn11-1262/f.2001.07.010>

- [20] Li, X., Wang, J., & Fu, H. (2002). An empirical study on the trading behavior of individual securities investors in China. *Economic Research Journal*, 47(11), 54–63+94. <https://doi.org/10.19744/j.cnki.11-1081/f.2002.11.006>
- [21] Wu, J., Wang, C., Chen, Z., et al. (2020). A study on the disposition effect of Chinese individual investors—A perspective of irrational beliefs. *Journal of Financial Research*, (2), 147–164.
- [22] Li, X., Yu, H., Lu, R., et al. (2011). An empirical study on behavioral biases of Chinese investors. *Management World*, (12), 1–12.
- [23] He, C., Chen, R., Xue, B., et al. (2021). Investor sentiment, limited arbitrage, and stock price anomalies. *Economic Research Journal*, (1), 180–197.
- [24] Li, T., Chen, H., Liu, W., et al. (2023). The role of social media sentiment in identifying irrational herding behavior in the stock market. *International Economic and Financial Review*, 87, 163–179.
- [25] Shleifer, A., & Vishny, R. W. (1997). The limits of arbitrage. *The Journal of Finance*, 52(1), 35–55. <https://doi.org/10.1111/j.1540-6261.1997.tb03817.x>

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Analysis of the Factors Influencing Residents' Cultural Consumption Expenditure——From the Perspective of Online Retail Use

Yulu Xia*

Faculty of Humanities and Social Sciences, University of Nottingham Ningbo China, Ningbo 315100, China

**Corresponding author: Yulu Xia.*

Abstract

On the basis of balanced provincial panel data from 30 provinces (autonomous regions and municipalities) in China for 2021 to 2024, this paper employs a two-way fixed effects model to empirically examine the impact of online retail development on residents' cultural consumption. It further explores the internal influencing mechanism through three pathways—consumption convenience, information accessibility and demand stimulation—and conducts robustness tests. The results indicate that the development of online retail has a significant positive effect on residents' cultural consumption. Similarly, the regional economic development level and residents' per capita disposable income also play positive driving roles in cultural consumption. The benchmark findings remain robust after alternative variable measurement methods, 1% winsorization, and one-period lag treatment of the core explanatory variables are adopted. Finally, targeted countermeasures and suggestions are proposed in terms of deepening industrial integration, improving logistics infrastructure, increasing residents' income, optimizing the supply of public cultural services, and promoting coordinated regional development. The research findings provide empirical support for tapping into the potential of residents' cultural consumption and facilitating the high-quality development of cultural consumption.

Keywords

online retail, residents' cultural consumption, two-way fixed effects, influencing factors, cultural consumption upgrading

1. Introduction

In the context of the rapid development of the digital economy and the continuous upgrading of residents' consumption structure, online retail has become a vital channel for expanding consumption and smoothing economic circulation, whereas cultural consumption serves as a core indicator for measuring residents' quality of life and social development level. With the continuous enrichment of online consumption scenarios, the role of online retail in reducing transaction costs, expanding consumption channels and improving consumption convenience has become increasingly prominent, with the influence on residents' cultural consumption increasing. Nevertheless, existing studies lack systematic analyses of the impact mechanism, regional differences and empirical evidence concerning the effect of online retail on cultural

consumption, and the quantitative correlation and transmission pathways between the two remain to be further clarified.

In view of the above, this paper adopts provincial panel data as research samples to empirically examine the impact of online retail development on residents' cultural consumption. It further conducts analyses from the perspectives of impact pathways and regional differences and performs robustness tests to ensure the reliability of the research conclusions. This study not only contributes to the research achievements in the fields of the digital economy and cultural consumption but also provides empirical references and policy implications for improving the quality and efficiency of online retail and promoting the expansion and upgrading of cultural consumption.

2. Literature Review

Early academic research on the platform economy and online retail is theoretically grounded in two-sided market theory, with a focus on analysing its transaction characteristics and the internal mechanism of efficiency improvement [4]. Relying on internet technologies, online retail breaks temporal and spatial constraints and features diverse transaction participants, digitalized transaction information, and platform-based transaction modes. As a core component of digital consumption, it drives consumption upgrades by expanding the coverage of supply and demand, reconstructing the connection between suppliers and consumers, and improving the efficiency of supply–demand matching [4]. The integration of digital and real economies, as the core form for the in-depth development of online retail, can reshape the consumption logic across the entire industrial chain of production, distribution, circulation and consumption. It promotes the quantitative and qualitative growth of consumption through precise demand matching, consumption capacity enhancement, and the liberation of time and space constraints on consumption [6]. The in-depth penetration of the digital economy further optimizes the consumption information environment, alleviates information asymmetry in the traditional consumer market, reshapes residents' consumption decision-making logic and demand preferences, and provides theoretical support for cultural consumption driven by online retail [10]. As the underlying foundation for the development of online retail, new digital infrastructure lays a hardware foundation for the popularization and upgrading of residents' cultural consumption by improving consumption convenience and broadening information access channels [11].

Cultural consumption is a core consumption type that meets residents' spiritual needs, and its influencing factors have long been a key research focus in academia. Through empirical analysis, Yuan Hongbo et al. (2013) reported that residents' personal interests, education level, age and occupation are the critical factors affecting cultural consumption. Moreover, income level and the accessibility of cultural facilities significantly affect residents' cultural consumption behaviors [1]. With the widespread popularization of digital technologies, the driving effect of digital channels on cultural consumption has become increasingly prominent. Improving residents' digital literacy and the application of digital intelligent platforms can effectively increase the accessibility of cultural consumption and increase cultural consumption by expanding income sources, strengthening information acquisition capabilities and optimizing social networks [2]. Digital inclusive finance relieves residents' liquidity constraints through digital services, improves their consumption capacity and willingness, and provides financial support for unlocking the potential of cultural consumption [12]. In addition, by innovating consumption modes and enriching consumption scenarios, digital technologies accelerate the optimization and upgrading of residents' consumption structure toward spiritual and cultural consumption, thereby further stimulating the cultural consumption market [13].

Extensive research has been conducted on the economic effects of online retail and the influencing factors of cultural consumption. Several studies have also discussed the moderating effect of digital intelligent platforms on the urban–rural cultural consumption gap. However, targeted research on the direct effect, internal mechanism and heterogeneous differences of online retail on residents' cultural consumption expenditure is lacking. Moreover, few studies have systematically analysed the specific paths through which online retail affects cultural consumption from the micro perspective. Accordingly, this paper adopts a two-way fixed effects model based on panel data to empirically test the impact of online retail development on residents' cultural consumption expenditure and further explores its internal mechanism and conducts heterogeneity analysis. These findings can not only expand the research perspectives on the influencing

factors of cultural consumption but also provide empirical evidence for revealing the potential of residents' cultural consumption.

3. Mechanism of Influence and Pathway Analysis

As the core form of digital consumption, online retail serves as a vital carrier for the popularization of online consumption scenarios. The popularization and development of online retail can affect residents' cultural consumption in three dimensions, namely, consumption cost, consumption scenarios and consumption willingness, and can act as a crucial driving force for the growth of cultural consumption. First, online retail increases cultural consumption by reducing consumption costs. Supported by internet technologies, online retail integrates the supply and demand information of cultural products with big data and digital logistics technologies, optimizes the allocation of cultural consumption resources, mitigates information asymmetry, and reduces intermediate costs such as offline search and transportation expenses. Therefore, residents can access cultural products and services at a lower cost. Second, online retail enhances cultural consumption options by expanding consumption scenarios. Leveraging digital technologies, it breaks the time and space constraints of traditional cultural consumption and gives rise to new cultural consumption modes, including online ticket booking, cultural and creative product online shopping, and digital content subscription. It broadens the boundary of cultural consumption from the supply side and stimulates residents' potential cultural consumption demand. Finally, online retail increases cultural consumption potential by increasing consumption willingness. Relying on platform-based operations, it provides accurate recommendations, word-of-mouth communication and convenient payments for cultural products; optimizes the cultural consumption experience; and accordingly increases residents' willingness and frequency of cultural consumption to drive the sustained growth of cultural consumption.

In the context of digital consumption, online retail has three prominent advantages. The first is the advantage of data factors. Online retail accumulates massive amounts of data, such as information on cultural consumption preferences and consumption behaviors, which serve as a core production factor driving the accurate matching of cultural consumption. The second is the advantage of digital technology. To improve the convenience and efficiency of cultural consumption, digital technologies, including the internet, logistics tracking and intelligent recommendation, should be applied in online retail. The third is the advantage of platform carriers. As the core carrier of online consumption, online retail platforms establish a new resource allocation mode for the docking of the supply and demand of cultural products. On the basis of the three major advantages of online retail, as well as its effects on cost reduction, scenario expansion and willingness improvement in terms of residents' cultural consumption, three pathways for online retail to promote the growth of residents' cultural consumption can be further identified, namely, the information accessibility pathway, the consumption convenience pathway, and the demand stimulation pathway.

Pathway One: Information Accessibility Pathway. Empowered by data factors, online retail improves the accessibility of cultural consumption information and thereby promotes the growth of residents' cultural consumption. Online retail platforms collect and integrate multidimensional information on the supply, price and evaluation of cultural products through big data technologies, converting scattered cultural consumption information into accessible and comparable effective elements. Such information not only helps residents quickly obtain cultural consumption-related content and eliminates the information isolation of offline cultural consumption but also enables them to accurately match their own cultural consumption preferences and reduce consumption decision-making costs. On the one hand, the improvement in information accessibility eliminates information barriers in cultural consumption, enabling more residents to reach more niche and high-quality cultural products. On the other hand, it strengthens residents' cognition and attention to cultural consumption, thus directly driving the increase in cultural consumption expenditure.

Pathway Two: Consumption Convenience Pathway. Empowered by digital technologies, online retail enhances the convenience of cultural consumption and further facilitates the growth of residents' cultural consumption. The in-depth application of digital technologies endows online retail with the features of borderless time and space, convenience and high efficiency. Residents can complete the selection, payment and after-sales service of cultural products anytime and anywhere, greatly reducing the time and physical costs of offline consumption. Moreover, the in-depth integration of online retail with logistics systems and mobile payments further improves the whole-process service of cultural consumption and particularly

provides equal cultural consumption opportunities for residents in rural and remote areas. Improved consumption convenience effectively relieves the restrictive conditions of residents' cultural consumption, transforms occasional cultural consumption into regular consumption, and promotes the simultaneous growth of the scale and frequency of residents' cultural consumption.

Pathway Three: Demand Stimulation Pathway. Empowered by platform carriers, online retail stimulates residents' potential cultural consumption demand and fuels the growth of cultural consumption. As important carriers of cultural consumption, online retail platforms break the scenario restrictions of traditional cultural consumption. Through new models such as live commerce, cultural and creative content marketing, and personalized recommendation, they create an immersive cultural consumption atmosphere and awaken residents' latent cultural consumption needs. In addition, the scale effect brought by platform operation lowers the prices of cultural products and further stimulates residents' cultural consumption willingness. Driven by demand stimulation, residents' cultural consumption demand shifts from passivity to initiative, and the consumption structure develops from singularity to diversification, ultimately achieving the steady growth of cultural consumption expenditure.

4. Research Design

4.1 Variable Selection

4.1.1 Explained Variable

Residents' cultural consumption (CC). In reference to the mainstream practices in the literature, the per capita collection volume of public libraries (volumes per capita) is adopted as the proxy variable for measurement. This indicator can directly reflect the supply level of regional cultural services and residents' cultural consumption capacity. With authoritative data sources and unified statistical caliber, it can effectively characterize the actual situation of regional residents' cultural consumption.

4.1.2 Core Explanatory Variable

Online retail development level (NR). The express delivery business volume (100 million pieces) is selected as a single proxy variable for measurement. This indicator is highly synergistic with the development of online retail and can intuitively reflect the market scale and degree of development of regional online retail. The data are sourced from the State Post Bureau and the National Bureau of Statistics of China. With official authority and strong continuity, it fully meets the measurement criteria for empirical research.

4.1.3 Mediator Variables

To verify the theoretical transmission mechanism proposed above, this paper selects information accessibility (Info), consumption convenience (Conve), and demand stimulation (consumption willingness, Willing) as mediator variables. Among them, information accessibility reflects the ability of online retail to reduce information barriers and optimize the information matching of cultural consumption; consumption convenience embodies the role of online retail in breaking temporal and spatial constraints and improving the convenience of cultural consumption; and demand stimulation indicates the capacity of online retail to awaken potential demand and enhance residents' cultural consumption willingness. In this paper, a mediation effect test is adopted to verify the transmission function of the above variables in the influence of online retail on residents' cultural consumption.

4.1.4 Control Variables

The intensity of digital consumption policy support (Digi) is measured as the proportion of fiscal expenditure related to the digital economy in general public budget expenditure. This indicator reflects the support of local governments for digital consumption, online retail and the digital transformation of cultural industries and has a vital effect on the development environment of online retail and the incentive intensity of cultural consumption.

The level of smart logistics infrastructure (Logi) is represented by a comprehensive index constructed by express outlet density and cold chain logistics coverage. As an essential foundation for the efficient operation

of online retail, smart logistics can improve the distribution efficiency of cultural products and increase residents' consumption convenience.

The supply density of cultural consumption scenarios (Cult) is measured by the number of new-type cultural consumption venues per 10,000 people. This indicator reflects the supply scale of immersive, digital and other emerging cultural consumption scenarios. It forms an online-offline complementary pattern with online retail and facilitates the increase in cultural consumption.

The dispersion of residents' consumption capacity (Capi) is measured by the Gini coefficient of residents' per capita consumption expenditure. It reflects the internal consumption gap within a region. An excessive consumption gap restrains the overall cultural consumption potential to a certain extent.

The degree of urban–rural consumption integration (Unif) is measured by the penetration ratio of online commodity consumption between urban and rural areas. This indicator reflects the balanced development of digital consumption in urban and rural regions. A higher level of urban–rural digital consumption integration helps expand the coverage of cultural consumption.

Residents' digital literacy level (Lite) is measured by the proportion of annual person-times of digital skill training in the permanent population. Improving residents' digital literacy can optimize the usage efficiency of online platforms, thereby increasing the adoption rate of online retail and residents' participation in cultural consumption.

4.2 Model Specification

4.2.1 Two-Way Fixed Effects Benchmark Model

Combined with the characteristics of panel data, this paper controls for provincial fixed effects and year fixed effects to alleviate endogeneity issues and constructs the benchmark model as follows:

$$CCit = \beta_0 + \beta_1 NRit + \sum \beta_k Controlit + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

As shown in Table 1, this paper defines and elaborates on all the variables in the model.

Table 1: Definition of Model Variables

Variable Symbol	Variable Name	Variable Definition/Description
i	Provincial dimension	Cross-sectional dimension of panel data, referring to 30 provincial-level regions across China (excluding Hong Kong, Macao, Taiwan and Xizang due to missing data)
t	Year dimension	Time dimension of panel data, denoting each year within the research period
CCit	Residents' cultural consumption	Cultural consumption level of province i in year t. As the explained variable of this paper, it measures the actual status of regional cultural consumption
NRit	Online retail development level	Online retail development level of province i in year t. As the core explanatory variable, it captures the influence of digital consumption channels
Controlit	Set of control variables	General term for control variables of province i in year t, including key factors such as digital consumption policy support and smart logistics infrastructure
μ_i	Provincial fixed effect	Controls the interference of time-invariant provincial inherent characteristics (e.g., geographical endowments, cultural traditions) on regression results
λ_t	Year fixed effect	Absorbs the common time-varying shocks such as macroeconomic conditions and policy changes on the explained variable
ε_{it}	Random disturbance term	Captures the impact of other omitted random factors and satisfies the basic assumptions of the classical linear regression model

4.2.2 Specific Functions of the Model

First, the core causal relationships should be accurately identified. The core research hypothesis of this paper is that the development of online retail promotes residents' cultural consumption. By simultaneously controlling for provincial fixed effects (μ_i) and year fixed effects (λ_t), the model can effectively eliminate the interference of region-specific time-invariant characteristics (such as cultural traditions and geographical endowments) and macro temporal shocks (including national digital consumption policies, the pandemic and other external events). This approach avoids misattributing the impact of irrelevant factors to online retail to

accurately capture the net effect of online retail on residents' cultural consumption and provide solid support for causal inference.

Second, the value of panel data should be fully considered. On the basis of provincial panel data, the model makes full use of dual-dimensional information, namely, cross-sectional differences across provinces and dynamic changes in a single province over time. Compared with pure cross-sectional data or time series data, it not only retains the richness of individual heterogeneity but also captures the dynamic evolution trend of variables, which greatly improves the efficiency and reliability of estimation results.

Third, endogeneity problems should be effectively alleviated. In view of the common omitted variable bias in empirical research (e.g., unobserved regional cultural atmosphere and national consumption trends), the model absorbs time-invariant and region-invariant omitted variables through fixed effects. It satisfies the unbiasedness assumption of the classical linear regression model, ensures the authenticity and effectiveness of coefficient estimation of the core explanatory variable, and provides a solid methodological guarantee for research conclusions.

This paper focuses on the latest integrated development trend of online retail and cultural consumption since 2021. Although the research period is relatively short with a limited time span, the research samples constitute balanced panel data. Moreover, the two-way fixed effects of province and year are controlled to absorb macroeconomic shocks to the greatest extent, rendering the empirical results of the model sound and referable.

4.3 Data Sources and Descriptive Statistics

4.3.1 Data Sources

In this paper, balanced provincial panel data from 30 provinces, autonomous regions and municipalities in China from 2021 to 2024 are selected as the research samples (the Xizang Autonomous Region is excluded because severe data are missing), with a total of 120 observations. Among the variables, residents' cultural consumption is measured by per capita public library collections, with data collected from the *China Culture and Tourism Statistical Yearbook* and annual statistical bulletins issued by the Ministry of Culture and Tourism. The development level of online retail is represented by express delivery business volume, and the relevant data are derived from postal and telecommunication statistics released by the State Post Bureau and the National Bureau of Statistics. The per capita regional gross domestic product and residents' per capita disposable income are sourced from provincial annual data of the National Bureau of Statistics and national economic and social development statistical bulletins of each province. All the data are officially released with high authority, complete integrity and no missing values, which fully meet the empirical analysis requirements of the two-way fixed effects model.

4.3.2 Descriptive Statistics

To clearly grasp the data distribution characteristics, dispersion degree and overall level of each variable, a descriptive statistical analysis is conducted on the core variables, and the results are presented in Table 2 below.

Table 2: Descriptive Statistics of the Major Variables

Variable Name	Symbol	Observations	Mean	Std. Dev.	Min	Max
Residents' Cultural Consumption	CC	124	0.95	0.55	0.42	3.40
Online Retail Development Level	NR	124	45.20	78.50	0.12	425.89
Per Capita Regional GDP	PGDP	124	85600.31	38200.47	42508.00	228167.00
Per Capita Disposable Income	Income	124	38500.62	15800.73	22066.00	88366.00

Note: The data in this table are collated and calculated on the basis of the panel data of 30 provincial-level regions from 2021 to 2024. The unit of residents' cultural consumption is volume per capita; the unit of the online retail development level is 100 million pieces; and per capita regional GDP and residents' per capita disposable income are measured in yuan. The descriptive statistics show that all the variables have a reasonable data distribution with no extreme outliers, which fully satisfies the data requirements for empirical research.

5. Empirical Analysis

5.1 Descriptive Statistics and Substantive Analysis

The descriptive statistical results indicate that the overall distribution of sample data is reasonable and can effectively support the results of the subsequent empirical analysis. There are obvious interprovincial gaps in residents' cultural consumption, with the overall mean being at a medium level, reflecting prominent imbalanced regional development. The large difference in the extremes of online retail development levels indicates a substantial disparity in the development of the digital economy and logistics industry across Chinese provinces. Specifically, the eastern coastal regions are far more advanced than the central, western and northeastern areas are. The dispersion of per capita regional GDP and residents' per capita disposable income is moderate, and the data fluctuation is within a reasonable range without extreme outliers. The overall data quality is reliable and can objectively reflect the actual economic development and residents' income level of each province. As shown in Table 3, the benchmark regression results verify the positive promotional effect of online retail.

Table 3: Descriptive Statistics

Variable Name	Symbol	Observations	Mean	Std. Dev.	Min	Max
Residents' Cultural Consumption	CC	120	0.95	0.55	0.42	3.40
Online Retail Development Level	NR	120	45.20	78.50	0.12	425.89
Per Capita Regional GDP	PGDP	120	85600.31	38200.47	42508.00	228167.00
Per Capita Disposable Income	Income	120	38500.62	15800.73	22066.00	88366.00

Note: The data in the table are compiled and calculated on the basis of the panel data of 30 provincial-level regions in China from 2021 to 2024. Residents' cultural consumption (CC) is measured by per capita public library collections (volumes per capita), and the online retail development level (NR) is measured by express delivery business volume (100 million pieces). The units of per capita GDP and residents' per capita disposable income are yuan. All the variables in this paper are measured by official single indicators with unified statistical caliber and authentic and reliable data.

5.2 Benchmark Regression and Substantive Analysis

In this paper, a two-way fixed effects model is adopted to empirically test the impact of online retail development on residents' cultural consumption. The Hausman test is conducted before regression to verify the rationality of adopting the fixed effect model. The overall regression results are statistically significant, with favourable model fitting performance, which can effectively explain the economic correlation between variables.

5.2.1 Substantive Analysis of Core Explanatory Variables

The regression coefficient of the online retail development level (NR) is significantly positive, indicating that the development of online retail has a prominent positive effect on residents' cultural consumption. In terms of mechanisms, online retail relies on digital technology to break the temporal and spatial constraints on cultural product consumption, greatly expanding access channels for cultural consumption. The continuous expansion of the express delivery scale further improves the circulation and distribution system of cultural products, reduces the transaction costs and access thresholds of residents' cultural consumption, effectively stimulates residents' willingness to consume cultural products, and ultimately increases the level of residents' cultural consumption. These results are highly consistent with the theoretical logic that the digital economy drives the upgrading of cultural consumption, which verifies the core research hypothesis of this paper.

5.2.2 Substantive Analysis of Control Variables

The regression coefficient of per capita GDP (PGDP) is significantly positive, suggesting that regional economic development is a crucial factor affecting residents' cultural consumption. Regions with greater economic development can invest more in public cultural facilities and provide better public cultural services, laying a solid material and hardware foundation for cultural consumption and thereby promoting the growth of cultural consumption. The coefficient of residents' per capita disposable income (Income) is also significantly positive, which conforms to the basic logic of consumption theory. As the foundation of consumption, higher disposable income strengthens residents' purchasing power for developmental consumption, such as cultural entertainment, increases the demand and willingness for public cultural consumption, and ultimately drives the continuous improvement of residents' cultural consumption. As

shown in Table 4, the benchmark regression results further confirm the positive promotion effect of online retail.

Table 4: Benchmark Regression Results

Variable	Two-way Fixed Effect Model
Online Retail Development Level (NR)	0.185***
Per Capita GDP (PGDP)	0.212***
Per Capita Disposable Income (Income)	0.168***
Constant	1.256***
Provincial Fixed Effect	Controlled
Time Fixed Effect	Controlled
R2	0.685
Observations	120

*Note: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. All the coefficients reported are standardized regression coefficients.*

5.3 Robustness Test and Substantive Analysis

To eliminate the interference of variable measurement bias, outliers and endogeneity on the benchmark regression conclusions and ensure the authenticity and reliability of the research results, this paper adopts three mainstream methods to conduct robustness tests, namely, replacing the measurement method of core variables, 1% winsorization treatment, and setting the core explanatory variable with a one-period lag. The test results are as follows.

5.3.1 Replacement of Core Variable Measurement Method

The core explanatory variable of the online retail development level (NR) is remeasured by replacing total express delivery volume with per capita express delivery volume to reestimate the model to examine the effects of different measurement indicators on the regression results. The results show that the regression coefficient of the online retail development level is still significantly positive and that the coefficient magnitude and significance are essentially consistent with those of the benchmark regression. These findings indicate that the research conclusions are not affected by alternative variable measurement methods.

5.3.2 1% Winsorization Treatment

To avoid interference from extreme data on the empirical results, all the continuous variables in the model are processed with 1% winsorization, and the regression is repeated after extreme outliers are excluded. The results indicate that the coefficient signs and significance of the core variables and control variables remain unchanged and that the model fitting effect does not obviously differ from that of the benchmark regression. These findings prove that extreme values do not distort the empirical conclusions of this study.

5.3.3 One-period Lag of the Core Explanatory Variable

To alleviate the reverse causality endogeneity between online retail and residents' cultural consumption, the online retail development level with a one-period lag (L.NR) is incorporated into the model for regression analysis. The results reveal that the coefficient of lagged online retail is still significantly positive, with the same influence direction as the benchmark regression. This reflects the sustained driving effect of online retail on cultural consumption and effectively relieves endogeneity bias. As shown in Table 5, the robustness test results are highly consistent with those of the benchmark regression.

Table 5: Robustness Test Regression Results

Variable	Benchmark Regression	Variable Replacement	1% Winsorization	One-period Lag
Online Retail Development Level	0.185***	0.183***	0.184***	0.181***
Per Capita GDP	0.212***	0.211***	0.211***	0.210***
Per Capita Disposable Income	0.168***	0.167***	0.167***	0.166***
Provincial Fixed Effects	Controlled	Controlled	Controlled	Controlled

Time Fixed Effects	Controlled	Controlled	Controlled	Controlled
R2	0.685	0.683	0.684	0.681
Observations	120	120	120	120

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Summary of Robustness Tests: The results of three robustness tests consistently demonstrate that the development level of online retail has a significant positive effect on residents' cultural consumption. The influence direction and significance of both the core and control variables are highly consistent with those in the benchmark regression. The benchmark regression conclusions are free from measurement bias, outlier interference and endogeneity problems, indicating that the research findings are authentic, robust and reliable.

5.4 Regional Heterogeneity Analysis

There are prominent differences among regions in China in terms of online retail infrastructure, residents' income levels and cultural consumption environments. To examine the regional differences in the impact of online retail on residents' cultural consumption, this paper divides the samples into three major regions—eastern, central and western China—in accordance with the official classification standards of the National Bureau of Statistics and conducts subsample regression.

The empirical results show that the coefficient of the online retail development level (NR) is significantly positive across all three regions, indicating that the promotional effect of online retail on cultural consumption is universally applicable. In terms of effect intensity, the eastern region has the largest coefficient, followed by the central region, while the western region has the lowest coefficient.

The main reason is that the digital economy in the eastern region is more mature with higher online retail penetration. Moreover, local residents have stronger willingness and purchasing power for cultural consumption, which together strengthen the driving effect of online retail. In contrast, restricted by inadequate digital infrastructure and relatively backwards economic development, the driving effect is relatively weak in central and western China.

The above findings verify that the impact of online retail on residents' cultural consumption presents significant regional heterogeneity. As shown in Table 6, the influence of online retail on cultural

Table 6: Regional Heterogeneity Regression Results

Variable	Eastern	Central	Western
Online Retail Development Level (NR)	0.215***	0.185***	0.152***
Per Capita GDP (PGDP)	0.236***	0.212***	0.187***
Per Capita Disposable Income (Income)	0.193***	0.168***	0.141***
Provincial Fixed Effects	Controlled	Controlled	Controlled
Time Fixed Effects	Controlled	Controlled	Controlled
R2	0.710	0.685	0.650
Observations	44	32	44

Note: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

5.5 Mechanism Test

To verify the theoretical mechanism proposed above, in this paper, information accessibility, consumption convenience and demand stimulation are selected as mediating variables, and a stepwise regression method is adopted to conduct mediating effect tests to verify the transmission paths through which online retail affects residents' cultural consumption.

First, this paper tests the total effect. The regression results show that the coefficient of online retail on residents' cultural consumption is significantly positive, indicating that online retail can significantly promote the improvement of residents' cultural consumption and that the total effect is valid.

Second, this paper examines the influence of the core explanatory variable on the mediating variables. The results indicate that online retail has significantly positive effects on information accessibility, consumption convenience and demand stimulation. These findings demonstrate that online retail can effectively increase the convenience for residents in terms of obtaining cultural consumption information,

optimize the overall level of cultural consumption convenience, and increase residents' potential cultural consumption demand.

Third, all the mediating variables are incorporated into the benchmark model for regression analysis. The results reveal that the coefficients of information accessibility, consumption convenience and demand stimulation are significantly positive, whereas the coefficient of online retail decreases correspondingly, which proves that the three variables all exert partial mediating effects.

The above findings indicate that online retail not only directly affects residents' cultural consumption but also indirectly promotes cultural consumption growth through three transmission paths: improving information accessibility, enhancing consumption convenience and stimulating potential consumption demand. The empirical results are highly consistent with the theoretical mechanism analysis in Chapter 3, which fully verifies the three proposed transmission paths. As shown in Table 7, the results of the mediating effect test confirm the validity of all three paths.

Table 7: Mediating Effect Test Results

Variable	Information Accessibility (Info)	Consumption Convenience (Conve)	Demand Stimulation (Willing)
Online Retail (NR)	0.178***	0.180***	0.176***
Control Variables (PGDP/Income)	Controlled	Controlled	Controlled
Provincial/Time Fixed Effects	Controlled	Controlled	Controlled
R2	0.678	0.680	0.676
Observations	120	120	120

*Note: *** indicates the 1% significance level, and the coefficients are standardized estimates.*

6. Research Conclusions

In this paper, balanced provincial panel data from 30 provinces (autonomous regions and municipalities) in China from 2021 to 2024 are adopted, and a two-way fixed-effects model is constructed to empirically test the impact of online retail development on residents' cultural consumption. Robustness tests are further carried out by replacing the variable measurement methods, conducting 1% winsorization, and lagging the core explanatory variable by one period. The main conclusions are summarized as follows:

6.1 The Development of Online Retail Has a Significant Positive Effect on Residents' Cultural Consumption

As measured by express delivery business volume, a higher level of online retail development can break the temporal and spatial restrictions on the circulation of cultural products, reduce consumption costs, and thereby effectively increase residents' cultural consumption level, which verifies the core research hypothesis of this paper.

6.2 Regional Economic Development and Residents' Income Serve as Essential Foundations for Driving Cultural Consumption

Both per capita GDP and per capita disposable income have significant positive effects on residents' cultural consumption. Regions with more developed economies and higher residential incomes have stronger cultural consumption capacity and greater utilization of public cultural services.

6.3 The Empirical Results of This Study are Robust and Reliable

The results of multiple robustness tests are consistent with those of the benchmark regression. There is no change in the influence direction or significance of the core variables, eliminating the interference of variable measurement, outliers, endogeneity and other factors on the conclusions. The research findings are authentic and valid.

7. Policy Recommendations

Combined with the empirical research conclusions of this paper, to further explore the driving effect of online retail on residents' cultural consumption, promote the steady growth of residents' cultural consumption level, and facilitate the high-quality development of cultural consumption, the following policy recommendations are proposed:

7.1 Deepening the Integrated Development of Online Retail and Cultural Consumption

E-commerce platforms, short video platforms and live streaming platforms should be encouraged to expand sales channels for cultural products and to increase the supply of cultural consumption, including books, cultural and creative products, and digital cultural content. Relying on digital technology to break the temporal and spatial constraints of cultural consumption reduces the circulation and transaction costs of cultural products and stimulates residents' cultural consumption potential through online retail innovation.

7.2 Improving Logistics Infrastructure to Ensure the Efficient Circulation of Cultural Products

Continuously optimize the layout of express delivery and logistics networks; increase investment in logistics infrastructure in central, western and rural areas; expand the coverage and distribution efficiency of express services; provide solid support for the cross-regional circulation of cultural products; and enable high-quality cultural resources to reach all residents more conveniently.

7.3 Raising Residents' Income to Consolidate the Foundation of Cultural Consumption

Efforts should be made to steadily increase residents' per capita disposable income by stabilizing employment, improving vocational skills and increasing income growth to increase residents' consumption capacity. Moreover, the social security system should be improved to reduce residents' precautionary savings and reduce their willingness to consume cultural entertainment, public cultural services and other services.

7.4 Optimizing the Supply of Public Cultural Services and Improving the Quality of Cultural Consumption

The investment in the construction of public facilities such as public libraries and cultural venues should be increased, public cultural service indicators, including per capita library collections, should be steadily improved, the equalization and universal accessibility of public cultural services should be promoted, and the conversion of public cultural resources into residents' actual cultural consumption should be improved.

7.5 Promotion of Coordinated Regional Development and Narrowing the Gap in Cultural Consumption

In view of the unbalanced development of cultural consumption and online retail among the eastern, central and western regions, interregional cooperation in the digital economy and cultural industries should be strengthened, digital platforms and logistics resources should be guided to favour underdeveloped areas, the regional disparity in cultural consumption levels should be gradually narrowed, and the coordinated development of cultural consumption should be realized.

References

- [1] Yuan, H. B., & Guo, F. L. (2013). Research on influencing factors of cultural consumption. *Science & Technology Vision*, (5), 33-34.
- [2] Yang, J. L., Li, X. X., Hu, Y. H., & Wu, L. P. (2024). Can digital literacy promotion promote rural residents' cultural consumption growth? *Journal of China Agricultural University*, 29(2), 261-272.
- [3] Wang, K. (2024, March 2). China's online consumption continues to improve. *People's Daily Overseas Edition*, 001.
- [4] Wang, G. S., & Wang, J. (2021). Economic analysis of online retail. *Journal of Beijing Technology and Business University (Social Sciences Edition)*, 36(1), 105-113.

- [5] Gu, J., Shu, H. J., & Tian, X. Y. (2024). Research on the influence mechanism of digital intelligent platforms in narrowing the urban–rural cultural consumption gap. *Rural Economy*, (5), 56-64.
- [6] Han, J., & Liu, Z. Y. (2024). Empowering and boosting residents' consumption through digital-real integration: Mechanism and promotion strategy. *Economic Review Journal*, (3), 60-69.
- [7] Qi, M. Y., & Ge, H. P. (2024). The impact of digital finance on rural consumption upgrading against the background of rural revitalization. *Productivity Research*, (3), 109-114.
- [8] Song, C., Zhong, J., & Li, M. (2024). The impact of coordinated development of logistics and rural e-commerce on rural residents' consumption upgrading. *Journal of Commercial Economics*, (5), 104-108.
- [9] Zhao, Y. M., & Yu, Q. Y. (2024). The impact of platform economy development on rural residents' consumption potential. *Journal of Commercial Economics*, (4), 42-46.
- [10] Zou, S. Z. (2024). The impact of digital economy development on residents' consumption preferences. *Modern Business*, (2), 46-48.
- [11] Yu, Z. P., & Chen, L. (2024). How do new digital infrastructure affect residents' green consumption? *Journal of Technical Economics & Management*, (1), 59-66.
- [12] Xie, H., & Liu, X. W. (2024). Research on the impact of digital inclusive finance on residents' consumption level. *Business Observation*, 12(1), 57-61.
- [13] Xiao, Z. L., Yang, Q., & Miao, Y. Z. (2024). Research on the role of digital technology in promoting residents' consumption upgrading. *Economist*, (1), 17-18.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

From Regulatory Text to Bank Behaviour: The Institutional Logic of the Basel NSFR Framework and Its Manifestation in HSBC

Chengji Liu*

University of Canterbury, 10 University Drive Christchurch 8041

**Corresponding author: Chengji Liu.*

Abstract

This paper examines the Basel III Net Stable Funding Ratio (NSFR) not simply as a technical liquidity requirement, but as a regulatory text that reflects an institutional logic of funding stability. Using qualitative policy text analysis, the study explores how the NSFR framework constructs behavioural expectations through its classifications, weighting system, and regulatory language, especially through the concepts of Available Stable Funding (ASF) and Required Stable Funding (RSF). To connect regulatory design with organisational response, the paper also uses HSBC as a case study and analyses its public disclosures on liquidity, funding, and balance sheet management. The findings suggest that the NSFR promotes a prudential banking model that prioritises stable liabilities, maturity alignment, and structural resilience. HSBC's disclosures largely reflect this logic in its funding strategy, risk governance, and liquidity reporting. The paper argues that the NSFR should be understood not only as a compliance ratio, but also as a governance mechanism that shapes legitimate banking behaviour.

Keywords

Net Stable Funding Ratio (NSFR), Basel III liquidity regulation, institutional logic

1. Introduction

The global financial crisis of 2008 exposed major shortcomings in the structure of bank financing. Many banks relied heavily on short-term wholesale financing. At the same time, they held long-term assets with poor liquidity. When market liquidity suddenly tightened, this mismatch made banks extremely vulnerable. This aggravated systemic instability (Basel Committee on Banking Supervision [BCBS], 2021) [1]. Basel III reforms were implemented by the Basel Committee in reaction to the crisis. Bank resilience was the goal of these efforts. Liquidity control is crucial among these reform initiatives. The committee developed two main tools. They are liquidity coverage ratio (LCR) and net stable funding ratio (NSFR). The former focusses on short-term liquidity pressure. While the latter focusses on the long-term stability of bank financing structures [1,2].

Compared with the LCR, the NSFR focuses on more structural issues. It requires banks to maintain sufficient and stable funding. It may adjust to their assets' liquidity and term structure and off-balance sheet exposures. NSFR intends to decrease banks' vulnerability to short-term funding. And it encourages banks to

establish a more sustainable financing structure within a one-year horizon [1]. This makes NSFR more than just a technical ratio. It is also a regulatory mechanism. It helps banks organize liabilities, manage assets, and strengthen the resilience of balance sheets.

Nevertheless, the majority of the current research considers the NSFR to be a prudential or technological need. Many studies have examined the impact of NSFR on profitability, lending, or compliance results, especially through quantitative analysis of bank-level data [3,4]. These studies are certainly valuable. However, less attention is paid to the regulatory text itself. In particular, they do not fully explore how the NSFR framework reflects a range of policy priorities and behavioural expectations through its language, categories, and weighting rules. This is an important research gap. Because regulation affects banks beyond formal restrictions. It defines prudential behavior and promotes particular strategic reactions.

In order to fill this research gap, this paper examines the NSFR from the perspective of institutional logic. Institutional logic refers to a broader system of values, assumptions and rules that guide organisational behaviour. As far as NSFR is concerned, its key logic lies in the stability of funding. It encourages banks to prioritize stable obligations to match long-term assets with stable cash. Also, it reduces structural dependency on short-term market funding. This logic is reflected not only in the overall policy objectives, but also in detailed classifications such as available stable funding (ASF) and required stable funding (RSF). These classifications give different weights to different types of liabilities and assets [1]. These classifications not only measure bank liquidity. They also show which financing methods are favoured and which are discouraged.

Therefore, this paper aims to explore how the institutional logic of funding stability is constructed in the framework of the NSFR of the Basel Accord. And how HSBC's strategic adjustment reflects this logic. To answer this question, this paper combines the qualitative text analysis of the Basel Committee documents and the case study of HSBC. HSBC is an ideal case. It is a globally active and systemically important bank. And it has a wide range of publicly disclosed information, including the annual report and the Pillar 3 report. This information enables this study to examine how the priorities contained in the NSFR framework are reflected in the bank's own discourse on financing strategies, balance sheet management and liquidity risk management [5,6].

This paper is expected to make three contributions. First of all, it enriches the literature in the fields of Basel III and liquidity regulation. It emphasizes that the NSFR is more than a technical regulatory tool. Instead, it should be regarded as a policy text with a unique institutional logic. Secondly, it contributes to the methodology. It combines regulatory text analysis with case-based interpretation of bank information disclosure. This allows this paper to link rule design with organisational response in a more interpretive way. Third, it provides an empirical perspective. It reveals how a large international bank openly explains its strategic adjustment in a language that is basically consistent with the logic of funding stability supervision.

2. Literature Review

2.1 Net Stable Funding Ratio and Basel III Liquidity Supervision

After the global financial crisis in 2008, liquidity regulation became the core content of Basel III. The Basel Committee introduced the LCR to cope with short-term pressure. And it introduced the NSFR to enhance the resilience of long-term structural funding. NSFR aims to reduce banks' dependence on unstable short-term financing and promote a better match between asset maturity and financing maturity [1,2]. Early academic studies also believe that NSFR is an important change in the field of prudential regulation. Because it directly addresses the problem of maturity transformation. And maturity transformation was a major weak link in the banking system during the crisis [7,8].

A large number of literature focusses on the economic impact of the NSFR. For example, King (2013) made one of the earliest detailed evaluations of NSFR [7]. He also pointed out that NSFR may improve the resilience of banks, but also affect their net interest margins. Research by Dietrich et al. (2014) shows that the business model of banks is crucial to their ability to meet new liquidity requirements [9]. This shows that the costs and benefits of NSFR vary between different institutions. Giordana and Schumacher (2017) also found that the Basel III standard, including liquidity rules, will change the behaviour and risk status of banks [10]. In

summary, these studies show that NSFR is not a neutral measurement tool. It has an important impact on the structure and performance of the banking sector.

At the same time, most literatures mainly regard the NSFR as a technical or prudential rule. The main question is usually whether the ratio can improve stability, affect profitability or change the lending capacity. Although this research idea is useful, it often pays less attention to the text of the regulations itself. In particular, it pays less attention to how the rule defines "stable" funding, and how its categories and weights guide the choices of banks [1,2].

2.2 Banks' Response to Liquidity Supervision

The second type of literature discusses how banks should respond to the liquidity and capital rules of Basel III. Some studies focus on balance sheet adjustment. For example, Wei et al. (2017) theoretically proved that the requirement of the NSFR will affect the bank's choice of debt maturity and asset structure [11]. Their research is of great significance. Because it shows that liquidity regulation can shape banking strategies by encouraging more stable liabilities and changing asset composition.

Other studies focus on the impact on performance. Le (2020) found that there is a nonlinear relationship between the NSFR and the profitability of banks [12]. This shows that a moderate increase in stable funding can improve efficiency. While excessive liquidity adjustment may reduce efficiency. Pak (2020) also shows that funding stability is crucial to profitability, especially when considering systemic importance [13]. These studies show that compliance with NSFR is not costless. It boosts bank risk resilience. But it will also bring about trade-offs in terms of profitability and business strategy.

Recent research has expanded this discussion to the field of lending and risk-taking. Veeramoothoo et al. (2022) found that the impact of Basel III liquidity regulation on the performance of U.S. banks varies according to the profitability of banks [14]. Sharma et al. (2023) believe that the liquidity and capital rules of Basel III can jointly affect lending, profitability and risk-taking [15]. They do not influence independently. Agnese et al. (2025) further pointed out that the NSFR constraint has an impact on the profitability and lending behaviour of EU banks [3]. In short, these documents show that banks do not simply "comply with" liquidity regulations. On the contrary, they will adjust the financing structure, loan behaviour and risk management according to regulatory pressure.

However, most of the existing literature is still result-oriented. They usually discuss the actions taken by banks after supervision, rather than how regulation itself defines the standard for prudential behaviour. In other words, the existing literature focusses more on behavioural effects than on the meaning structure of the rules themselves. And this is where the perspective based on the text and the system can play a role.

2.3 Institutional Logic and Regulatory Discourse

In order to make up for this research gap, this paper draws on the perspective of institutional logic. Institutional logic includes assumptions, values, and regulations. Organisations perceive situations and take decisions based on these systems [16,17]. Later evaluation studies found that institutional logic well explains how the system influences organizational stability and change [18](Wu et al., 2023).

This study will use this perspective. Because NSFR is more than a statistic. ASF and RSF classifications assume which obligations are more reliable and which assets need more solid support. The rule assesses more than liquidity risk. The ideal banking behavior model is also defined. Institutional logic links regulatory design with organizational reaction.

The study of discourse and institutional change also supports this view. Brown et al. (2012) believe that institutional change is often promoted through rhetorical techniques in key texts [19]. A recent study by Zilber (2024) shows that institutional logic is not only an abstract concept [20]. It can also be realised through narrative language, so as to establish coherence and legitimacy. These studies show that the policy text defines what is rational, legitimate and ideal. Applied to banking supervision, this means that the text of the Basel Accord can not only be studied as a legal rule. It can also be studied as a governance text that conveys priorities and expected goals.

2.4 Research Gaps and Contributions

The existing literature has laid a solid foundation, but there are still gaps in research. The research on the NSFR mainly focusses on its technical design and its impact on profitability, loans and financial stability [3,7,9,12]. The study on the response of banks to prudential supervision mainly focusses on the measurable adjustment of financing structure, asset allocation and risk-taking behaviour [11,14,15]. At the same time, institutional logic literature shows that text, categories and classification rules can shape organisational behaviour by defining what is legal and appropriate behaviour [16,17,21]. However, this view is rarely applied to Basel liquidity regulation, especially the NSFR.

This paper aims to fill this research gap and explore how the NSFR framework of the Basel Accord builds the institutional logic of financing stability. And explore how this logic is reflected in HSBC's strategic disclosure. In this way, this paper shifts the focus of discussion from NSFR as merely a technical compliance ratio to understanding it as a regulatory text. This contains policy priorities and encourages specific banking behaviour patterns [2,22]. This method is a supplement to the existing empirical research. And it provides a more interpretive way. This links the rule design of the Basel Accords with the strategic adjustment at the banking level.

3. Methodology

3.1 Research Problems and Method Selection

The problems of this study include two interrelated parts. The first part discusses how the NSFR is constructed as a regulatory text. How does it define stable funding? How does it classify assets and liabilities? And what behaviours does it encourage. The second part discusses how this logic is reflected in the public disclosure of a large international bank. Since this study focusses on meaning, classification and behavioural expectations, qualitative research design is the most appropriate choice [23,24].

More specifically, this paper adopts policy text analysis as the main method. And this paper takes the case study of HSBC as an example. Policy text analysis helps to explain how the Basel Committee conveys regulatory priorities through definitions, weights, and normative language. The case of HSBC provides a practical situation. It is used to examine whether and how this regulatory logic is reflected in the bank's strategy and information disclosure. The case study is not used to derive a broad statistical generalisation. It provides an example of analytical value. This example shows how regulatory logic may become organizational practice [25].

3.2 Using Policy Text Analysis as the Main Method

The core approach of this study is to conduct a qualitative document analysis of the Basel Committee texts. Literature analysis is widely applied in qualitative research. The policy document not only lists the rules. Moreover, it also expresses the definitions of assumptions, priorities and appropriate actions [23]. This is particularly important for NSFR. Because this framework is built on ASF and RSF classifications. It also explains flexibility, prudent financing and structural liquidity risk [1].

This paper not only regards the NSFR as a technical ratio. It also regards the NSFR as a policy text that carries the logic of a specific system. From the perspective of institutional logic, rules are not just formal constraints. They also reflect values and shape the organization's perception of legitimate and rational behaviour [16,17]. Therefore, the analysis of this paper focusses on the three dimensions of the NSFR framework:

- (i) How to define and prioritise funding stability
- (ii) How ASF and RSF classifications reflect behavioural preferences
- (iii) How can normative language and technical language convey regulatory expectations?

The purpose of this paper is not to count the number of words mechanically. Instead, it explains how the regulatory categories and language build a coherent understanding of prudential banking behaviour. This

interpretive method is consistent with the previous research results. These studies show that policy texts and institutional discourse play an important role in shaping organisational behaviour and legitimacy [21,24].

3.3 Data Sources

The main data of text analysis comes from the official documents of the Basel Committee. In particular, documents directly related to Basel III liquidity reform and the NSFR. These documents include the implementation and evaluation report of the Basel Committee, which elaborates on the purpose, structure and expected effect of the reform [1,2]. These documents not only show the formal regulatory design, but also explain the broader policy considerations behind it.

This case study uses the public information of HSBC. This includes annual reports, strategic reports and Pillar 3 disclosure documents. These documents show how HSBC presents its financing situation, liquidity management, balance sheet strategy and risk governance to regulators and investors [5,6]. Because this study explores how the regulatory logic is reflected in the strategic adjustment. Therefore, these disclosing documents are an appropriate source of evidence.

There are three reasons for choosing HSBC Bank as a case for analysis. First of all, HSBC is a large-scale bank with global business. It has long been regarded as a systemically important organisation in the global banking industry. This makes it an ideal case for studying the impact of the Basel III liquidity standard. Secondly, HSBC has extensive and detailed public disclosure information, especially in its annual report and Pillar 3 report. This makes document-based analysis possible. Third, HSBC's business model integrates retail banking, commercial banking and global market business. This enables this study to observe how the logic of structural funding is reflected in different areas of banking strategy.

Taking HSBC as an example, it is not intended to show that all banks have the same response. On the contrary, the case aims to clarify how the institutional logic explained in the text of the Basel Accord can be translated into specific organisational practices. This is in line with the logic of case studies, that is, the role of cases is to deepen analytical understanding, rather than to pursue statistical representativeness [25].

3.4 Analysis Procedure

3.4.1 Text Analysis of the Basel Accord

At this stage, this paper will carefully study the Basel Committee document to identify key concepts, classifications and recurring topics. Focus on the terms such as “stable funding”, “resilience”, “maturity profile” and “structural liquidity risk”, as well as the logic of ASF and RSF weights. The purpose is to clarify how the NSFR can build a prudential banking model. And what types of financing and asset structures it tends to favour.

3.4.2 HSBC Information Disclosure Analysis

At this stage, this paper will examine the public documents of HSBC. This study will identify the alignment with the institutional logic stipulated in the text of the Basel Accord in terms of wording and strategic discussion. The analysis focusses on three aspects: financing structure, asset and liability management and liquidity/risk governance. The purpose is to examine whether HSBC's information disclosure also emphasises financing stability, maturity matching and structural resilience.

This two-stage design ensures that the case analysis is always closely related to the main research issues. The paper does not assume from the beginning that HSBC's change is simply due to the adjustment of the NSFR. On the contrary, the paper first clarifies the logic of the regulatory text. Then examines whether similar priorities have emerged in the case materials.

3.5 Scope and Limitation

The advantages of this method are obvious. But at the same time, there are some limitations. Its main advantage is that it can link regulatory design with organisational response in an interpretive and theoretical way. It is especially suitable for revealing that banking supervision not only works through rigid constraints. It also plays a role through various categories, expectations and legitimising language.

At the same time, this study does not try to measure the causal effect of the NSFR on all banks. This study also does not claim that HSBC can represent all global systemically important banks (G-SIBs). This research relies on public documents, which reflect strategic communication and organisational practice. Therefore, the research results should be understood as analytical and interpretive. The results of this study are not statistically universal.

4. The Institutional Logic of the Framework of the Net Stable Funding Ratio

4.1 NSFR as a Rule to Measure “Stable” Banking Business

The Basel Committee regards the NSFR as a core component of the post-crisis Basel III reform. According to the BCBS (2014) [22], the NSFR requires banks to maintain a stable funding profile that matches their on- and off-balance-sheet business composition. The aim is to reduce the possibility of interruptions in the source of regular funds leading to a deterioration of bank liquidity and an increase in the risk of bankruptcy. This expression proves NSFR is more than a technical ratio. It also shows regulatory inclinations. Thus, banks should use a stable, predictable, and market-resistant funding structure.

This language expresses institutional reasoning to protect finances. Therefore, the NSFR is more than a liquidity metric. The definition of “good” banking is also promoted. Resilience and structural robustness are prioritized over short-term financial flexibility and opportunistic wholesale market dependency. The follow-up assessment of the Basel Committee also confirmed this broader goal [2]. BCBS explained that liquidity reforms, including NSFR, are aimed at improving the resilience of the banking industry. It also aims to reduce the transmission of financial shocks to the real economy. The recent monitoring of the Basel Committee also shows that the NSFR level of large international active banks remains stable. This shows that the standard has become a regular component of prudential regulation [26].

4.2 How ASF Factors Generate Priority Liabilities

This logic is reflected in the ASF framework of the NSFR. The standard allocates different ASF coefficients for different liabilities and capital instruments. Regulatory capital and liabilities with a residual maturity of one year or more obtain a 100% ASF coefficient. “Stable” retail and small business deposits with a term of less than one year get a 95% ASF coefficient. And “less stable” retail deposits get 90% of the ASF coefficient. In contrast, some short-term non-retail financing projects only get 50% of the ASF coefficient. Some liabilities that do not belong to the priority category receive an ASF coefficient of 0% [22]. Funding sources are sorted instead of being equal in importance. Long-term financing and stable deposits will be rewarded. Under prudential supervision, mobile and unpredictable funds are considered less ideal.

This rating is not neutral. It has a behavior preference. The NSFR does not require banks to raise retail deposits or long-term debt. Instead, it gives certain liabilities a higher value within the regulatory framework. In this way, the rule plays the role of an indirect governance mechanism. Through its classification and weight, it encourages banks to build a prudent financing structure that regulators consider appropriate. In other words, NSFR is governed through classification. It transforms the overall goal of stability into a set of practical hierarchies of preferred liabilities [16,22].

4.3 How RSF Factors Create Priority Assets

A similar logic is also reflected in the “required stable funding” (RSF) coefficient on the asset side. According to the Basel Accord, cash, central bank reserves and some very short-term central bank claims are subject to an RSF coefficient of 0%. The RSF coefficient of 5% applies to unencumbered Level 1 assets. And some short-term secured lending transactions attract a 10% or 15% RSF coefficient. In contrast, the RSF coefficient of non-high-quality assets, long-term loans, residential mortgages, other performing loans, stocks, commodities and certain encumbered or derivatives-related assets is much higher, usually between 50% and 100% [22]. In short, the less liquid, the longer the maturity or the higher the structural requirements of the asset, the more stable funds the bank needs to hold for it.

This structure gives rise to a second institutional logic: maturity alignment. The NSFR assumes that assets with longer maturities and poorer liquidity should be supported by more stable and longer-term liabilities. This

assumption directly challenges the financing model before the crisis. Before the crisis, banks relied heavily on short-term wholesale funding to support their long-term positions. Therefore, NSFR not only identifies liquidity risks after the fact, but also defines the form that a better balance sheet should have. It encourages banks to reduce structural maturity mismatches. And it regards financing options as part of risk management, rather than just cost-based financial decision-making [7,22].

4.4 Technical Rules as a Form of Behaviour Guidance

A prominent feature of the NSFR is that its behavioural message is conveyed through highly technical language. The standard is constructed with the concepts of categories, maturity, percentage and collateral handling. However, this technical form does not make its value neutral. On the contrary, it gives normative selection a scientific and objective appearance. For example, allocate 95% or 90% of asset stabilisation funds (ASF) to retail deposits. And only 50% or 0% are allocated to certain short-term non-retail liabilities. This reflects the judgement of which sources of funds are more reliable under pressure. Similarly, the higher RSF coefficient of long-term and non-high-quality liquid assets (non-HQLA) also reflects this view. That is, these assets need stronger structural support [22].

This means that the mechanism of action of the NSFR is guided by behaviour rather than direct instruction. It does not stipulate every treasury decision. Instead, it changes the incentive environment for balance sheet management. Therefore, it is very useful to analyse NSFR as a policy text. The rule is technical in form, but its effect is normative. It informs banks which financing methods are prudential financing, which liabilities are prioritised, and which asset structures need to be more strictly constrained. According to institutional logic, NSFR is a regulatory document that sets values and guides organizational behavior [17,21].

5. HSBC Case: How is the Regulatory Logic Reflected in Practice?

5.1 Why HSBC is a Useful Verification Case

HSBC is an ideal case for studying how this regulatory logic is reflected in practice. Because it is a large-scale bank with business all over the world. And it has made detailed public disclosures in terms of liquidity, financing and risk governance. In its 2024 annual report and accounts, HSBC pointed out that the average liquidity and financing positions of all its major operating entities are higher than the minimum level required by the regulatory authorities. And each entity has maintained a sufficient and stable financing situation. This situation is evaluated by the NSFR or other appropriate indicators [5]. This statement is highly consistent with the statement on stability, adequacy and structural financing discipline in the Basel Accord.

5.2 Evidence of the Logic of “Stability First” in HSBC’s Financing Model

The clearest evidence is reflected in the NSFR data disclosed by HSBC. HSBC reported that it manages financing risks in accordance with the NSFR rules of the Prudential Regulation Authority (PRA), and the average NSFR in 2024 is 143%. As of June 30, 2024 and December 31, 2023, the ratio is 138% [5]. HSBC also reported that by the end of 2024, its available stable funding is US\$1,523 billion, and the required stable funding is US\$1,064 billion. These data show that HSBC not only meets the standards, but also has a considerable buffer space above the minimum requirement of 100%. This is in line with the logic of “stability first”. Funding stability is not considered a secondary compliance issue. Instead, it is regarded as the core goal of the balance sheet.

HSBC’s liability structure also helps to explain this result. In the 2024 annual report, the total amount of customer accounts was US\$1,654.955 billion, while in 2023 it was US\$1,611.647 billion, and the total deposits were US\$1,728.952 billion [5]. In the Pillar 3 disclosures of the group, the total amount of customer accounts in the regulatory balance sheet is US\$1,680.704 billion. The issuance of debt securities is US\$104.434 billion [6]. The NSFR gives special preferential Available Stable Funding factor (ASF) treatment for stable retail and small business deposits. And long-term liabilities are given full ASF treatment [22]. Therefore, HSBC’s huge deposit base is very suitable for the liability-side preferences built into NSFR. In this sense, the bank's financing structure seems to be roughly consistent with the regulator's preference for stable deposit financing rather than more fragile short-term market financing.

5.3 Evidence of Term Matching in HSBC's Balance Sheet Management

The balance sheet disclosed by HSBC also reflects the term matching logic of the NSFR. In its 2024 annual report, HSBC explained that part of the growth of NSFR in 2024 was offset by the increase in required stable funding, which mainly came from the increase in financial investments and derivatives activities [5]. It directly reflects the structure of the Basel Accord. When banks increase their business activities in areas that require more stable structural funding, regulatory indicators will respond by increasing the ratio of RSF. Therefore, HSBC's explanation shows that the bank understands and uses categories similar to NSFR itself to present its balance sheet.

More broadly, HSBC said that it has adopted a wide range of liquidity and financing measures, and each major operating entity maintains sufficient unrestricted current assets and a stable financing situation [5]. This way of expression is very important. It shows that liquidity management is not limited to maintaining short-term survival under pressure. This also includes long-term structural balance sheet management. This view is consistent with the basic logic of the NSFR. The rule encourages banks to regard assets and liabilities as interrelated through long-term financing discipline. HSBC's disclosure wording also reflects the same perspective [5,22].

5.4 Behavioural Guidance Evidence in Governance and Disclosure

The third area where regulatory logic can be reflected is corporate governance. HSBC pointed out in its Risk Review 2024 that its risk management framework supports the identification, evaluation, management and reporting of risks in a consistent manner, and the clarification of responsibility. The report also points out that the Group's risk and compliance department provides challenge, oversight and appropriate balances in risk/return decision-making. At the same time, it helps employees' behaviour and risk preferences to be consistent [27]. These statements do not refer to the NSFR in isolation, but show how prudential regulatory expectations can be integrated into the bank's internal governance system. Regulatory logic has become a part of management reasoning. It is not just a part of external compliance.

The disclosure of the Pillar 3 report further reinforces this point. HSBC clearly links its liquidity report to regulatory requirements. And it points out that it has strengthened the liquidity integration process in 2024 and revised the relevant provisions previously aimed at addressing historical limitations. The bank also said that the strengthened process also affected the reporting of the LCR and the NSFR [6]. This is a powerful example of behaviour guidance. Once the rules are established, banks will not only adjust their financing structure, but also adjust their internal systems, reporting processes and disclosure structures around regulatory categories. In this sense, NSFR does not only constrain HSBC from the outside. Instead, it becomes integrated into the way banks monitor themselves and present the soundness of their balance sheets.

5.5 Summary

The third area where regulatory logic can be reflected is corporate governance. HSBC pointed out in its Risk Review 2024 that its risk management framework supports the identification, evaluation, management and reporting of risks in a consistent manner, and the clarification of responsibility. The report also points out that the Group's risk and compliance department provides challenge, oversight and appropriate balances in risk/return decision-making. At the same time, it helps employees' behaviour and risk preferences to be consistent [27]. These statements do not refer to the NSFR in isolation, but show how prudential regulatory expectations can be integrated into the bank's internal governance system. Regulatory logic has become a part of management reasoning. It is not just a part of external compliance. As summarised in Table 1, this pattern is also consistent with the wider links between NSFR institutional logic and HSBC's disclosed funding structure, balance-sheet management, and governance practices.

Table 1. From NSFR Institutional Logic to HSBC's Strategic Response

NSFR Organisational Logic	The way it is presented in the text of the Basel Agreement	How does HSBC realise this?

Stability-first logic	NSFR requires a “stable funding profile” and a “sustainable funding structure” to reduce the risk of failure[22].	HSBC said that all important entities maintained sufficient stable funding and reported that the average net stable funding ratio in 2024 was 143%, higher than the regulatory minimum standard [5].
Prefer to stabilise liabilities	Stable retail and small business deposits can get 95% ASF; less stable retail deposits can get 90% ASF; some short-term non-retail funding sources can only get 50% or 0% ASF [22].	HSBC has several deposits. The annual report estimates 1.655 trillion US dollars in client accounts. The regulatory balance sheet shows 1.681 trillion US dollars in client accounts [5,6].
Maturity alignment logic	The RSF coefficient of long-term, poor liquidity and non-HQLA assets is relatively high, usually as high as 85% or 100% [22].	HSBC explained that higher financial investment and derivatives activities have increased the demand for stable funds. This indicates that its own report follows the same structure as the NSFR logic [5].
Behaviour guidance through governance	NSFR operates through classification, weighting and technical rules instead of direct operational commands [22].	HSBC links liquidity management, risk appetite, reporting and internal process improvement with prudential supervision and regulatory disclosure [6,27].

The HSBC case validates this paper’s key points. The NSFR goes beyond a technical ratio certified at the conclusion of the reporting period. This has apparent institutional logic. It tends to favour stable liabilities, closer term matching, and stronger liquidity and financing risk management. HSBC’s statement does not indicate NSFR drives all strategic modifications. Business models, market conditions, and domestic regulatory expectations also matter. However, the case does show that HSBC’s public statements, indicators and reporting focuses are basically consistent with the logic contained in the text of the Basel Accord. Therefore, the case of HSBC can very well verify the text-centred analysis method of this paper [5,6,22].

6. Conclusion

In general, the analysis of this study shows that NSFR should not only be regarded as a technical liquidity rule. It should be regarded as a regulatory text aimed at promoting a specific prudential banking model. In the Basel Accord, funding stability is regarded as an ideal and reasonable cornerstone of bank resilience. And it classifies liabilities and assets through the ASF factor and the RSF factor. This gives this logic practical effectiveness. In other words, NSFR does not only measure the balance sheet status. It more actively defines the form that a better financing structure should have.

Therefore, the main contribution of this paper lies at the theoretical level. Most of the existing literature regards the NSFR as a prudential regulatory indicator. And they focus on issues such as compliance, profitability or lending effects. This paper starts from different perspectives and points out that NSFR also contains institutional logic. Its rules reflect the banking industry’s concept of putting stability first, its preference for maturity matching, and the indirect behaviour guidance achieved through classification and weighting. It shifts the focus from the mechanical effect of the rules to how the rules define prudent and legitimate banking behaviour. This helps to link banking regulation with broader institutional and organisational theories.

The second contribution of this paper is at the methodological level. It does not rely solely on quantitative indicators. Instead, it takes policy text analysis as the main method, and takes HSBC as an example for explanation. Through a detailed interpretation of the Basel Committee document, this paper reveals how the language and classification of the NSFR construct a regulatory logic. This paper explains how this approach is applied in the actual organization by evaluating HSBC’s public disclosure information. This technique is useful. It connects how banks respond to regulations with how those regulations are designed. This is something that most traditional compliance research often ignores.

The case of HSBC clearly supports this argument. HSBC’s disclosure information repeatedly uses words such as “sufficient stable funds”, “structural liquidity strength” and “stable balance sheet management”. With its massive deposit base and explanation that financial investment and derivatives trading demand dependable

finance, its average NSFR in 2024 is 143%. All of them demonstrate the bank's Basel Framework-compliant reporting rationale. HSBC does not base every strategy decision purely on NSFR. Many variables will impact banks. This encompasses macroeconomics, corporate models, domestic regulation, and market expectations. Even yet, the instance indicates that a big global bank's funding and liquidity discussions and management are strongly rooted in Basel Accord funding stability rationale.

These findings also have broader implications for regulatory research. They show that the role of prudential regulation is not only reflected in rigid requirements and numerical thresholds. Its role is also reflected in the classification, definition and legitimising language. This is especially evident in the NSFR. The rule does not clearly stipulate which liabilities should be issued or what assets should be held by the bank. Instead, it has established a hierarchical structure of prioritised sources of financing. And it has built a disciplined asset-liability structure. In this sense, NSFR realises regulation by shaping the acceptable range of choices. For researchers, this means that the regulatory text itself should be regarded as an independent governance mechanism, not just as a background rule in the empirical model.

At the same time, this paper also has obvious limitations. It only focusses on the case of HSBC. Therefore, it cannot be claimed that its conclusion is statistically universal for all G-SIBs. In addition, this paper relies on publicly disclosed information, which is partly a strategic communication document. This may favour the image of the bank. In addition, this paper does not try to distinguish the NSFR from other Basel III reforms (such as the LCR or capital requirements). These limitations should be clearly pointed out. However, they do not affect the main purpose of this paper. This paper aims to show how regulatory logic is constructed in the text and how this logic is reflected in the strategy disclosed by large banks.

References

- [1] Basel Committee on Banking Supervision. (2021). A report to G20 leaders on implementation of the Basel III regulatory reforms. Bank for International Settlements. <https://www.bis.org/bcbs/publ/d510.htm>
- [2] Basel Committee on Banking Supervision. (2022). Evaluation of the impact and efficacy of the Basel III reforms. Bank for International Settlements. <https://www.bis.org/bcbs/publ/d544.htm>
- [3] Agnese, P., Delle Foglie, A., Porretta, P., & Santoboni, F. (2025). How does the NSFR regulatory constraint affect profitability and lending? Evidence from EU banks. *European Financial Management*. <https://doi.org/10.1111/eufm.70032>
- [4] Sidhu, A. V., Rawal, A., & Goyal, V. (2022). Net stable funding ratio (NSFR) and bank performance: A study of Indian banks. *Journal of Risk and Financial Management*, 15(11), 527. <https://doi.org/10.3390/jrfm15110527>
- [5] HSBC Holdings plc. (2025a). Annual report and accounts 2024. <https://www.hsbc.com/-/files/hsbc/investors/hsbc-results/2024/annual/pdfs/hsbc-holdings-plc/250219-annual-report-and-accounts-2024.pdf>
- [6] HSBC Holdings plc. (2025b). Pillar 3 disclosures at 31 December 2024. <https://www.hsbc.com/-/files/hsbc/investors/hsbc-results/2024/annual/pdfs/hsbc-holdings-plc/250219-pillar-3-disclosures-at-31-december-2024.pdf>
- [7] King, M. R. (2013). The Basel III net stable funding ratio and bank net interest margins. *Journal of Banking & Finance*, 37(11), 4144–4156. <https://doi.org/10.1016/j.jbankfin.2013.07.017>
- [8] Gobat, J., Yanase, M., & Maloney, J. (2014). The net stable funding ratio: Impact and issues for consideration (IMF Working Paper No. 14/106). International Monetary Fund. <https://www.imf.org/external/pubs/ft/wp/2014/wp14106.pdf>
- [9] Dietrich, A., Hess, K., & Wanzenried, G. (2014). The good and bad news about the new liquidity rules of Basel III in Western European countries. *Journal of Banking & Finance*, 44, 13–25.

<https://doi.org/10.1016/j.jbankfin.2014.03.041>

- [10] Giordana, G. A., & Schumacher, I. (2017). An empirical study on the impact of Basel III standards on banks' default risk: The case of Luxembourg. *Journal of Risk and Financial Management*, 10(2), 8. <https://doi.org/10.3390/jrfm10020008>
- [11] Wei, X., Gong, Y., & Wu, H.-M. (2017). The impacts of Net Stable Funding Ratio requirement on banks' choices of debt maturity and asset structures. *Journal of Banking & Finance*, 82, 229–243. <https://doi.org/10.1016/j.jbankfin.2017.02.006>
- [12] Le, M., Hoang, V.-N., Wilson, C., & Managi, S. (2020). Net stable funding ratio and profit efficiency of commercial banks in the US. *Economic Analysis and Policy*, 67, 55–66. <https://doi.org/10.1016/j.eap.2020.05.008>
- [13] Pak, O. (2020). Bank profitability in the Eurasian Economic Union: Do funding liquidity and systemic importance matter? *The North American Journal of Economics and Finance*, 54, 101265. <https://doi.org/10.1016/j.najef.2020.101265>
- [14] Veeramoothoo, S., & Hammoudeh, S. (2022). Impact of Basel III liquidity regulations on U.S. bank performance in different conditional profitability spectrums. *The North American Journal of Economics and Finance*, 63, 101826. <https://doi.org/10.1016/j.najef.2022.101826>
- [15] Sharma, A. K., & Chauhan, R. (2023). Impact of Basel III liquidity and capital regulations on bank lending and financial stability: Evidence from emerging countries. *Journal of Corporate Accounting & Finance*, 34(4), 28–45. <https://doi.org/10.1002/jcaf.22630>
- [16] Thornton, P. H., & Ocasio, W. (2008). Institutional logics. In R. Greenwood, C. Oliver, K. Sahlin, & R. Suddaby (Eds.), *The SAGE handbook of organizational institutionalism* (pp. 99–129). SAGE. <https://doi.org/10.4135/9781849200387.n4>
- [17] Thornton, P. H., Ocasio, W., & Lounsbury, M. (2012). *The institutional logics perspective: A new approach to culture, structure, and process*. Oxford University Press. <https://global.oup.com/academic/product/the-institutional-logics-perspective-9780199601936>
- [18] Wu, X., Tan, X., & Wang, X. (2023). The institutional logics perspective in management and organizational studies. *Journal of Business Research*, 167, Article 114183. <https://doi.org/10.1016/j.jbusres.2023.114183>
- [19] Brown, A. D., Ainsworth, S., & Grant, D. (2012). The rhetoric of institutional change. *Organization Studies*, 33(3), 297–321. <https://doi.org/10.1177/0170840611435598>
- [20] Zilber, T. B. (2024). Narrating institutional logics into effect: Coherence across cognitive, political, and emotional elements. *Administrative Science Quarterly*, 69(1), 172–221. <https://doi.org/10.1177/00018392231217712>
- [21] Zilber, T. B. (2017). How institutional logics matter: A bottom-up exploration. In J. Gehman, M. Lounsbury, & R. Greenwood (Eds.), *How institutions matter!* (Research in the Sociology of Organizations, Vol. 48A, pp. 137–155). Emerald Publishing Limited. <https://doi.org/10.1108/S0733-558X201600048A005>
- [22] Basel Committee on Banking Supervision. (2014). Basel III: The net stable funding ratio. Bank for International Settlements. <https://www.bis.org/bcbs/publ/d295.pdf>
- [23] Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40. <https://doi.org/10.3316/QRJ0902027>

- [24] Tracy, S. J. (2019). *Qualitative research methods: Collecting evidence, crafting analysis, communicating impact* (2nd ed.). Wiley-Blackwell. <https://www.wiley.com/en-be/Qualitative+Research+Methods%3A+Collecting+Evidence%2C+Crafting+Analysis%2C+Communicating+Impact%2C+2nd+Edition-p-9781119390800>
- [25] Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE. <https://uk.sagepub.com/en-gb/eur/case-study-research-and-applications/book250150>
- [26] Basel Committee on Banking Supervision. (2025). Basel III monitoring report. Bank for International Settlements. <https://www.bis.org/bcbs/publ/d592.pdf>
- [27] HSBC Holdings plc. (2025c). Risk review 2024. <https://www.hsbc.com/-/files/hsbc/investors/hsbc-results/2024/annual/pdfs/hsbc-holdings-plc/25027-risk-review-2024.pdf>

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open - access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Correlation Analysis of Time Value of Cash Flow and Pricing in IPO Enterprises —— Based on Multi-period Cash Flow Model

Hongwei Huang*

Accounting College, Southwestern University of Finance and Economics, Chengdu 611130, China

**Corresponding author: Hongwei Huang.*

Abstract

In recent years, growth and technology-based companies account for a relatively high proportion of A-share IPO enterprises, and the fluctuation of cash flow and rising capital expenditure make the issue pricing more uncertain. In this paper, based on the multi-period cash flow model, the correlation between the time value of cash flow and IPO pricing is discussed, taking the IPO enterprises of science and technology innovation board Growth Enterprise Market as samples from 2024 to 2025. Through sample selection, unified parameter setting, correlation analysis, unary linear regression and comparison of pricing deviation, this paper finds that the direct explanatory power of discount rate to IPO issue price is weak, and the explanatory effect of DCF model is limited in the whole sample, but it is obviously enhanced after excluding abnormal samples, while PE method is closer to the actual issue price as a whole. The research shows that the multi-period cash flow model has a conditional explanatory function in IPO pricing, and its applicability is constrained by the stability of cash flow and the market pricing mechanism.

Keywords

initial public offering, time value of cash flow, DCF valuation, PE method

1. Introduction

In 2023, China's capital market realized the registration system for the main board, science and technology innovation board, Growth Enterprise Market and North Stock Exchange. The market pricing of IPO gradually shifted from administrative guidance to market-oriented inquiry mechanism, and the pricing and rationality of IPO gradually became the core issue of capital market reform. However, some IPO companies appear pricing deviation after listing, which leads to scientific thinking about IPO pricing logic.

The time value of cash flow is the core cornerstone of modern financial asset pricing theory. As a classic method of enterprise valuation, multi-period cash flow model (DCF) is an important theoretical basis for IPO pricing, which is to discount the future cash flow of an enterprise to the present according to the time value of money. In the actual IPO pricing process, multiple factors, such as the issuance mechanism, market expectation, comparison of the same industry and investor sentiment, will affect the final issue price, and the setting of the discount rate itself has strong parameter sensitivity and subjectivity [1].

The existing academic research on IPO pricing and cash flow valuation has achieved rich results. In recent years, domestic academic research shows that Song Shunlin thinks that the institutional background of IPO pricing in China is different from that in the United States, so there are many differences between IPO pricing and the United States, and there are many theories such as information asymmetry theory [2]. Recent studies have shown that the IPO issue price is not a direct mapping of the intrinsic value of the enterprise, but a market-oriented result formed under the joint action of bookkeeping, investor concern and underwriter incentive [3-5]. At the same time, IPO pricing deviation and quotation correction will also be affected by liquidity, social media attention and underwriter ranking [6-8]. At the valuation level, DCF model is still an important tool to identify the intrinsic value of enterprises, but its results are highly sensitive to the setting of discount rate, cash flow stability and risk handling methods [9-11]. Therefore, around the recent A-share IPO enterprises, there is still room for further analysis to test whether the time value of cash flow can be reflected in the actual issue price, and the differences between the multi-period cash flow model and PE method. Based on this, this paper takes A-share IPO enterprises from 2024 to 2025 as samples, investigates the correlation between the time value of cash flow and IPO pricing, and compares the explanatory effects of DCF model and PE method.

Based on the above theoretical analysis, this paper puts forward the following assumptions.

H1: There is a correlation between the discount rate of IPO enterprises and the IPO issue price, but its explanatory power may be obviously restricted by market pricing mechanism and investors' expectations.

H2: There is a significant correlation between the intrinsic value of enterprises calculated by multi-period cash flow model and the actual IPO issue price.

2. Research Design

2.1 Sample Selection and Data Source

This paper takes the A-share IPO enterprises from 2024 to 2025 as the research object, and the sample range is the companies that completed IPO and went public in the A-share market from 2024 to 2025. Key data such as issue price, EPS, P/E ratio, cash flow from operating activities and capital expenditure can be obtained from the prospectus, financial statements and issuance announcement of sample enterprises, while the samples of enterprises with missing relevant data and inconsistent financial caliber are excluded. Finally, after screening, this paper finally retained 50 IPO companies listed in science and technology innovation board and Growth Enterprise Market as research samples.

The data used in this paper come from: first, the prospectus, listing announcement and annual financial statements of enterprises; Second, the relevant financial indicators and data disclosed in the cash flow statement, balance sheet and income statement of the enterprise; Thirdly, the market-level data adopts risk-free interest rate ERP and China ten-year government bond yield, and the market risk premium is uniformly set at 6%; Fourth, the industry comparability coefficient is reflected by β coefficient; Fifth, the company's prospectus, listing announcement and annual financial statements all come from Juchao Information Network, and the company's cash flow statement, balance sheet and income statement come from Oriental Fortune Network.

In order to ensure the uniformity of caliber among samples, some parameters are set uniformly in this paper: the market risk premium is 6% and the long-term growth rate is 3%.

2.2 Definition of Variables

The explained variable in this paper is the IPO issue price per share, which is recorded as P. The IPO issue price can directly reflect the market pricing of enterprises in the process of listing, and it is also the core object of this paper.

The core explanatory variables in this paper are as follows. First, the discount rate is recorded as R. In this paper, the simplified CAPM method is used to estimate the discount rate, and its calculation formula is:

$$r_i = R_f + \beta_i \times ERP \quad (1)$$

Among them:

r_i represents the discount rate of the i th enterprise and R_f represents the risk-free interest rate. In this paper, 1.82%, β_i represents the β coefficient of the i th enterprise, and ERP represents the market risk premium. In this paper, 6% is uniformly taken.

Second, the intrinsic value of DCF per share is recorded as V_{DCF} . This paper first calculates the free cash flow, then predicts the cash flow in the next five years, and predicts the final value, and finally gets the total value of the enterprise and the intrinsic value per share.

In order to analyze the IPO issue price more comprehensively, this paper sets the following control variables:

PE p/e ratio This paper adopts the issue price-earnings ratio, and further calculates the PE valuation method. Second, the enterprise scale. In this paper, the total assets are taken as the enterprise scale index. Third, ROE. This paper uses the return on net assets to measure the profitability of enterprises. The stronger the profitability, the easier it is to support higher issuance pricing in theory.

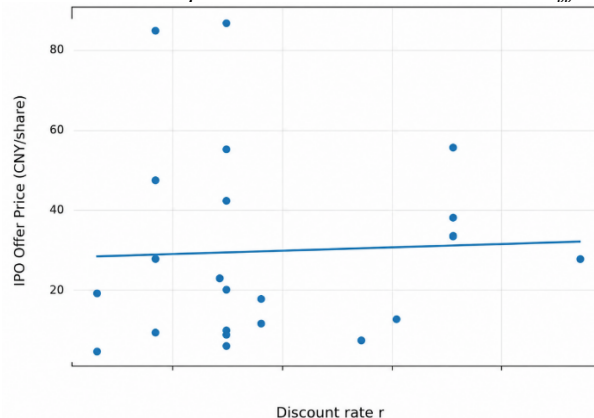
The core parameters of DCF valuation are uniformly set as follows. The long-term growth rate is 3%, which reflects the growth level of the enterprise entering a stable period and avoids excessive amplification of the final value, which is both reasonable and feasible. The risk-free interest rate is 1.82%, and the yield of China's 10-year government bonds in the sample period is adopted to meet the discounted cash flow demand in the next five years. The market risk premium is 6%, which adapts to the higher risk characteristics of IPO enterprises and conforms to the general valuation interval, which can reduce short-term fluctuation interference and improve sample comparability.

3. Practice Test Results and Analysis

Hypothesis 1 holds that there is a significant correlation between the discount rate of IPO enterprises and the IPO issue price, and it is negatively correlated as a whole. In order to test this hypothesis, this paper makes correlation analysis and unitary regression analysis on the discount rate and IPO issue price per share of 50 sample enterprises.

The results of correlation analysis (as shown in Figure 1) show that the correlation coefficient between the discount rate and the IPO issue price is -0.0265, and the results of regression analysis show that the regression coefficient is -50.2716, but the determining coefficient R^2 is only 0.0007, which shows that the explanatory power of the discount rate to the IPO issue price is very weak. The core reason for this result is that the formation mechanism of IPO issue price of A shares is more driven by market factors, and it is difficult to directly map the time value of funds and risk compensation logic represented by discount rate to the actual issue price. The IPO issue price itself has strong market-oriented pricing characteristics, and market expectation and growth premium also occupy an important position in IPO pricing. At the same time, the differentiation of discount rate in the sample is limited, which also weakens its statistical interpretation ability. This result is consistent with the conclusion of IPO pricing research in recent years. Relevant research shows that the issue price is formed under the joint action of bookkeeping, investors' attention and price discovery mechanism, and market sentiment and issuance mechanism will weaken the direct explanatory power of single valuation parameters such as discount rate to the issue price [4-6]. Therefore, assumption 1 is not supported.

Figure 1: Relationship between discount rate and IPO offer price



Hypothesis 2 holds that there is a significant correlation between the intrinsic value of enterprises calculated by multi-period cash flow model and the actual issue price of IPO. This paper also uses correlation analysis and regression analysis to test this hypothesis.

The whole sample test results show that the correlation coefficient between DCF intrinsic value per share and IPO issue price is -0.0614, the regression coefficient is -0.0436, and the determining coefficient R^2 is 0.0038, as shown in Figure 2. It shows that the relationship between them is not significant under the condition of full sample. However, after excluding the samples with negative DCF intrinsic value, the correlation coefficient between DCF intrinsic value per share and IPO issue price rose to 0.5351, the regression coefficient was 0.5484, and the determining coefficient R^2 was 0.2864, as shown in Figure 3, which showed an obvious positive correlation. Considering the particularity of IPO enterprises, DCF model has strong explanatory power in normal samples, which shows that DCF model has certain advantages in stabilizing the valuation of enterprises. Therefore, assumption 2 is partially supported. Judging from the valuation research in recent years, this result is reasonable. Relevant literature points out that although DCF model has the function of intrinsic value identification in a highly uncertain environment, its results are highly sensitive to discount rate setting, cash flow stability and risk treatment methods, so it is more prone to valuation bias in samples with large operating fluctuations, and its explanatory power is usually stronger in samples with stable cash flow [12-14].

Figure 2: DCF intrinsic value and IPO offer price (full sample)

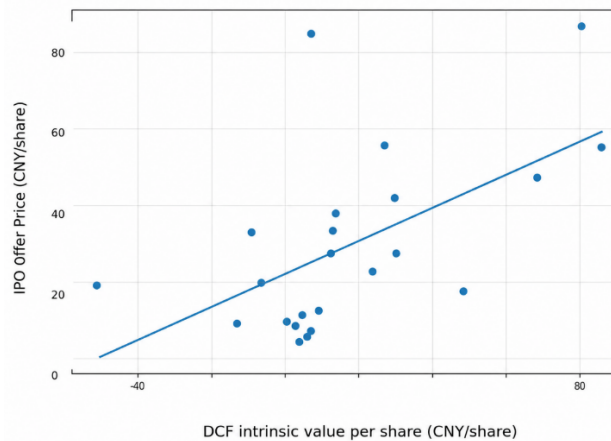
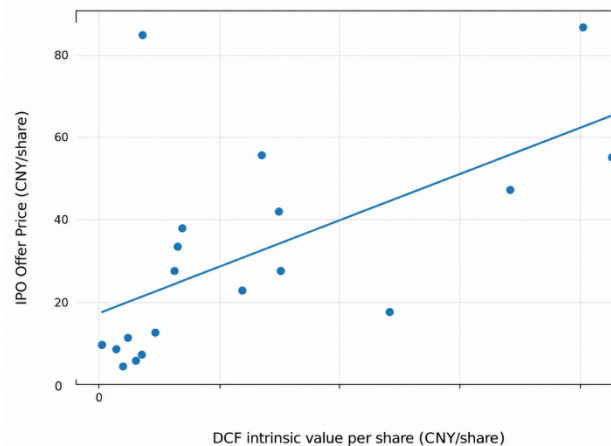


Figure 3: DCF intrinsic value and IPO offer price (excluding abnormal samples)



4. Result Analysis

This paper draws the following conclusions through three hypothesis tests. First, there is a weak negative correlation between the discount rate and the IPO issue price as a whole, and the linear relationship is not significant, indicating that the time value has limited influence on pricing. This is the result of multiple factors,

such as cash flow characteristics of enterprises, market sentiment, issuance mechanism and underwriter's reputation, and is not simply determined by the discount rate. Secondly, the intrinsic value of DCF has no obvious explanatory power to the issue price under the whole sample; After excluding the samples with negative cash flow, the two showed a significant positive correlation. It shows that the applicability of DCF depends on the stability of enterprise cash flow, which is easy to be distorted for IPO enterprises with large cash flow fluctuations in the growth period and has strong explanatory power for enterprises with stable operation. Thirdly, PE method is superior to DCF model as a whole, which is highly consistent with the IPO pricing practice in China, and also consistent with the related research on IPO pricing deviation and quotation correction in recent years [8-10]. In practice, pricing needs comprehensive quotation correction, market acceptance, liquidity and investor structure, and the relative valuation method is more realistic. DCF is superior in theory, but it is highly sensitive to the discount rate, growth rate and free cash flow, and it is easy to produce a big deviation under the unstable cash flow and unified parameter setting of IPO enterprises.

From the market level, IPO pricing is not only determined by intrinsic value, but also influenced by investor preference and market sentiment. At the enterprise level, the samples are mostly in the growth stage, with high capital expenditure, unclear profit model and more stable PE method; At the institutional level, China's inquiry mechanism strengthens the logic of relative valuation, and PE method is more in line with the current system and market habits.

5. Research Conclusions and Suggestions

This paper takes A-share IPO enterprises from 2024 to 2025 as the research object, based on the multi-period cash flow model, investigates the relationship between the time value of cash flow and IPO pricing, and compares the explanatory effects of DCF model and PE method. The results show that the direct explanation of the discount rate to the IPO issue price is weak, indicating that the actual issue pricing is more influenced by market valuation, industry comparison and investor expectations; The explanatory effect of DCF intrinsic value per share in the whole sample is limited, but the correlation is obviously enhanced after excluding abnormal samples, which shows that the multi-period cash flow model has certain applicability in enterprises with stable cash flow. In contrast, PE method is closer to the actual issue price as a whole, showing stronger explanatory power of realistic pricing. Based on the above conclusions, combined with the research results, this paper puts forward suggestions from three aspects: for issuers, the PE method should be the core of IPO pricing, and the DCF model should be used as an auxiliary reference for intrinsic value. DCF can reflect the long-term value, but it is easily influenced by the fluctuation of parameters and cash flow. PE method is more suitable for the pricing of a shares. The issuer can use DCF to determine the value range first, and then use PE method to price according to the industry valuation and market environment to improve the pricing rationality. For investors, it is necessary to recognize the limitations of DCF in IPO valuation and be wary of high deviation risks, especially for companies with growth and large cash flow fluctuations. When investing, we should combine profit stability, cash flow quality and capital expenditure structure. For stable enterprises, we can refer to DCF, and for high-growth enterprises, we should make a comprehensive judgment with PE method and industry valuation to enhance the robustness of decision-making. For regulators, we should strengthen the disclosure of IPO valuation information and improve pricing transparency. The issuer is required to simultaneously disclose DCF and PE valuations and their deviations from the issue price, reveal risks and differences, improve pricing efficiency, curb irrational high valuations, and promote the standardized and rational development of the IPO market. There are still some limitations in this paper. First, the limited number of samples and the short sample time interval may affect the universality of the conclusion; Second, the parameters are set uniformly, which fails to fully reflect the individual differences of enterprises; Third, the empirical analysis method is relatively basic, and more complicated grouping and robustness test have not been further developed. In the future, the sample range can be expanded, the research interval can be extended, and the explanatory power and academic value of the research can be further enhanced by combining industry, development stage and multiple regression analysis.

References

- [1] Wei, L., & Yu, R. H. (2024). Creative research on enterprise IPO valuation under the registration-based system. *Management & Technology of SMEs*, (24), 176–178.

- [2] Song, S. L. (2022). Chinese-style IPO pricing: A literature review. *Journal of Central University of Finance & Economics*, (1), 51–63.
- [3] Sun, F., Yin, C., Zhou, S., & Zhu, Z. (2022). IPO underpricing and mutual fund allocation: New evidence from registration system. *International Review of Financial Analysis*, 84, 102405.
- [4] Hiraki, T., & Ito, A. (2023). Two-step price adjustments of IPO book building in Japan. *Pacific-Basin Finance Journal*, 78, 101977.
- [5] Liu, L. X., Lu, R., Sherman, A. E., & Zhang, Y. (2023). IPO underpricing and limited attention: Theory and evidence. *Journal of Banking & Finance*, 154, 106932.
- [6] Espenlaub, S., Mohamed, A., & Saadouni, B. (2024). Underwriter incentives and IPO pricing. *Journal of Corporate Finance*, 87, 102625.
- [7] Bhurjee K, Saragur Ramanna V. Preopening auctions and price discovery in initial public offerings[J]. *Journal of Banking & Finance*, 2024, 165: 107196. DOI: 10.1016/j.jbankfin.2024.107196.
- [8] Rossovski J, Lucey B, Helbing P. Determinants of IPO overpricing[J]. *British Journal of Management*, 2025, 36(1): 383-399. DOI: 10.1111/1467-8551.12858.
- [9] Farooque O A, Chowdhury A, Tarique K M. The impact of lead underwriter ranking on the offer price revision of IPO firms: Evidence from the US[J]. *Investment Analysts Journal*, 2025. DOI: 10.1080/10293523.2025.2578128.
- [10] Han Y, Zhang C. Social media analysts, investor attention, and IPO underpricing[J]. *The British Accounting Review*, 2025: 101748. DOI: 10.1016/j.bar.2025.101748.
- [11] Guo Y, Li L Y, Zhong H. Bigger pie, bigger slice: Liquidity, value gain, and underpricing in IPOs[J]. *Journal of Financial Markets*, 2025, 72: 100949. DOI: 10.1016/j.finmar.2024.100949.
- [12] Huang S, Tan H, Wang X, Yu C. Valuation uncertainty and analysts' use of DCF models[J]. *Review of Accounting Studies*, 2023, 28(2): 827-861. DOI: 10.1007/s11142-021-09658-w.
- [13] Lof M, Nyberg H. Discount rates and cash flows: A local projection approach[J]. *Journal of Banking & Finance*, 2024, 162: 107127. DOI: 10.1016/j.jbankfin.2024.107127.
- [14] Vonlanthen, J. (2024). On the determinants of discount rates in discounted cash flow valuations: A counterfactual analysis. *Real Estate*, 1(2), 174-197.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

The Impact of Artificial Intelligence Industry Development on Service Exports: A Systematic Review Approach

Youyu Pan*

School of Business, Shanghai Jian Qiao University, Shanghai 202309, China

**Corresponding author: Youyu Pan.*

Abstract

With the rapid development of the digital economy, artificial intelligence (AI) has become an important factor influencing international service trade. As the growth of traditional merchandise trade slows, digitally deliverable services have continued to expand, supported by the increasing role of data and AI technologies. Using a systematic literature review approach and drawing on recent macro-level trade statistics and micro-level empirical studies, this paper examines the mechanisms through which AI affects the scale and structure of service exports and reviews the related empirical evidence. Recent data show that digitally deliverable services account for more than half of global service trade. Existing studies suggest that AI can improve the technological sophistication of knowledge-intensive service exports by lowering cross-border search costs, increasing total factor productivity, and enhancing service tradability. In this sense, AI contributes not only to the expansion of service trade but also to its movement toward higher value-added activities. The findings of this study help explain recent developments in digital trade and provide policy implications for narrowing the global digital divide, particularly through improvements in digital infrastructure and increased investment in research and development.

Keywords

artificial intelligence, service exports, digitally deliverable services

1. Introduction

1.1 Contemporary Context and Development Status of Digital Service Exports

Amid increasing uncertainty in the global economic recovery, the structure of international trade has continued to change. Supported by advances in artificial intelligence (AI), digitally deliverable services have become an increasingly important source of global trade growth. The impact of AI on service exports can be better understood in the context of recent macroeconomic trade data.

In 2023, while global merchandise trade contracted overall, digital service exports achieved counter-cyclical growth, expanding by 9% year-on-year to reach USD 4.25 trillion. From 2005 to 2023, global digitally deliverable service trade grew at a compound annual growth rate (CAGR) of 8.2%, significantly outpacing merchandise trade (4.6%) and overall service trade (6.2%) over the same period. By 2023, digital service

exports accounted for 13.8% of total global exports of goods and services, and more than half (54.2%) of total service exports. These figures indicate that digitally delivered services now occupy an increasingly important position in global trade.

However, significant regional disparities remain in both AI commercialization and digital service exports. In 2023, Europe maintained the largest share of digital service exports (52.4%, approximately USD 1.62 trillion), while Asia recorded the strongest growth, rising to 23.8% (approximately USD 1.15 trillion). North America accounted for 17.5%. Overall, developed economies monopolized over 80% of the global digitally deliverable service trade market. A key prerequisite for AI to effectively empower service exports is the presence of robust digital infrastructure. According to the World Bank's 2024 report [1], median internet connection speeds in high-income and upper-middle-income countries increased by 50% in 2023–2024, reaching 143 Mbps. In contrast, median speeds in low-income countries remained stagnant below 25 Mbps. Differences in computing infrastructure and network capacity continue to constrain the adoption of generative AI technologies in many developing countries, which may further widen international productivity gaps.

Despite these disparities, China has placed increasing policy emphasis on the development of digital trade. In recent years, China has hosted the China International Fair for Trade in Services (CIFTIS) at a consistently high level. This reflects both the scale of China's domestic market and the growing importance of service trade in the country's opening-up strategy.

1.2 Practical Challenges: Examining Service Industry Bottlenecks Through the Lens of Trade Deficits

In stark contrast to the persistent trade surplus in manufacturing, China's service trade has long faced a structural deficit. This deficit is closely related to the relatively limited scale, efficiency, and standardization of China's service sector.

$$NX = Y - (C + I + G) \quad (1)$$

From the perspective of the macroeconomic national income trade identity (Equation 1), net exports essentially represent the difference between China's total output (productive capacity) and total domestic demand (absorption). China's manufacturing sector maintains strong export competitiveness largely because of its high production efficiency and large-scale industrial capacity. In contrast, the service sector is more dependent on human capital and customized provision, which limits its ability to achieve large-scale output expansion and export growth [2]. Therefore, improving the productive capacity of the service sector remains important for expanding service exports.

1.3 Technological Breakthrough: AI-Empowered “capacity Revolution” in the Service Sector

The growing adoption of artificial intelligence may help alleviate some of the scale constraints traditionally associated with service industries. Unlike earlier internet technologies, In particular, generative AI (AIGC) has increased the degree of standardization in certain service activities. By automating repetitive cognitive tasks, AI enables some traditionally customized services to be delivered in more standardized and digitalized forms. Owing to its large domestic market and extensive data resources, China provides a favorable environment for the application and development of AI technologies in the service sector. With AI technological empowerment, the service sector is expected to achieve a “capacity leap” similar to that of manufacturing—substantially increasing output density and quality without proportionally increasing human capital investment. This may strengthen the sector's ability to meet domestic demand while also supporting the expansion of service exports.

1.4 Research Focus and Logical Framework of This Paper

Although the academic community has reached a preliminary consensus on the trade-promoting effects of artificial intelligence, existing literature predominantly focuses on macro-level effects. However, existing studies have paid relatively limited attention to the mechanisms through which AI affects service sector capacity and trade patterns [3]. In light of this, this study combines recent trade data with existing literature to examine the mechanisms by which AI influences service exports. The logical structure of this paper is as follows: First, it systematically reviews and evaluates representative literature on the trade effects of AI.

Second, it focuses on constructing a theoretical mechanism framework for AI's role in enhancing service capacity and reducing transaction costs. Finally, it summarizes the research conclusions and puts forward targeted policy recommendations grounded in China's reality as a super-large market.

2. Literature Review

2.1 International Research: Digital Economy Theoretical Frameworks and Global Macro Effects

Early international studies examined the effects of digital technologies on trade from both theoretical and empirical perspectives. Goldfarb and Tucker [3] discussed how algorithms and digital technologies influence global economic activities by reducing information processing costs. At the empirical level, Aghion et al. [4] examined the potential long-term growth-driving effects of artificial intelligence, finding that AI adoption increases the degree of automation in service production processes in high-income countries. Recent studies have increasingly focused on trade barriers associated with digital technologies. Reports from the Organisation for Economic Co-operation and Development (OECD) indicate that between 2014 and 2023, global trade barriers on digitally enabled services increased by 25%, primarily due to intensified government regulation of cross-border data flows and artificial intelligence algorithms [5].

2.2 Chinese Research: Improvements in Export Technological Complexity and Domestic Empirical Evidence

In recent years, Chinese scholars have increasingly examined the role of artificial intelligence in improving China's position in global value chains through service trade, providing robust support from micro-level panel data. Regarding the enhancement of technological complexity, Yao Tingting and Bao Yanan [6], using the artificial intelligence development index published by Stanford University and panel data from 29 major economies, conducted empirical tests. Their results indicate that AI contributes to improvements in the technological complexity of service exports, primarily through two channels: improving service production efficiency and optimizing the human capital structure of the service industry. Further, in exploring the moderating role of R&D and innovation, Tao Mengyao [7], based on cross-country panel data, demonstrated that the positive effect of AI on the technological complexity of service exports is highly robust. Moreover, a country's overall "innovation level" exerts a significant positive moderating effect in this process. This finding suggests that the adoption of AI technology alone may not be sufficient to generate competitive advantages; it must be accompanied by high-intensity R&D investment to translate AI capabilities into stronger competitiveness in knowledge-intensive service exports.

3. Theoretical Mechanisms of Ai-Empowered Service Exports From the Perspective of the Macroeconomic Trade Identity

Existing studies have mainly examined the trade effects of digital technologies from the perspective of firm-level transaction costs. However, this perspective provides limited explanation for China's persistent service trade deficit at the macro level and the growing role of artificial intelligence in addressing this issue [3]. Accordingly, this paper employs the national income trade identity $NX=Y-(C+I+G)$ as the analytical framework, where NX represents net exports, Y denotes China's total output (i.e., productive capacity), and $(C+I+G)$ represents total domestic absorption (i.e., domestic demand). From a macroeconomic perspective, AI may promote service exports by changing the production structure of the service sector. This may increase domestic service capacity and support the expansion of service exports into international markets. The specific theoretical mechanisms can be decomposed into the following three dimensions:

3.1 Overcoming "Baumol's Cost Disease": AI-Driven Standardization and Reshaping of Service Sector Capacity

Within the framework of the macroeconomic trade identity, the absolute scale of total service output (Y) constitutes the cornerstone of export potential. The deep-rooted logic behind China's long-term service trade deficit lies in the rigid physical constraints on traditional service sector capacity expansion — a phenomenon

known in economics as “Baumol’s cost disease.” The production function of traditional knowledge-intensive services (including legal consulting, medical diagnosis, language translation, etc.) typically exhibits strict linear characteristics: $Y=f(L)$, where output scale is highly dependent on the time and effort invested by professional service personnel (L). Since service delivery often requires “co-location and simultaneous presence of producer and consumer in time and space” (e.g., face-to-face consultations), production efficiency is extremely difficult to improve significantly through capital equipment investment alone. This characteristic prevents traditional services from adopting the assembly-line production models common in manufacturing, resulting in limited capacity scale, high marginal costs, and an inability to effectively meet the increasingly diverse service consumption demands of China’s super-large market. As a result, domestic service supply has often been insufficient to meet demand, contributing to continued reliance on imported services.

To explain it, the following equation is introduced.

$$Y = A \cdot f(K_{Data}, L_{Expert}) \quad (2)$$

The rapid development of artificial intelligence — particularly generative AI and large language models — has altered traditional service production processes, driving its evolution toward a nonlinear structure (Equation 2), where A represents the level of technological progress and K_{Data} denotes computing power and data capital inputs.

This transformation operates through three core mechanisms.

First, structural substitution of production factors has enabled a transition from “human-driven” to “algorithm-driven” modes. Relying on natural language processing and machine learning algorithms, AI achieves large-scale replacement of high-density, repetitive mental labor in traditional service sectors. As highlighted by Acemoglu and Restrepo’s (2018) “race between man and machine” theory [8], automation technologies not only substitute for humans in routine tasks but also reconstruct the composition of production factors. Recent empirical research by Brynjolfsson et al. (2023) on generative AI further confirms [9] that AI significantly boosts labor productivity in the service sector, particularly by accelerating the standardization of tacit knowledge. This evolutionary process, in which “computing capital” offsets and substitutes for “human capital,” liberates professional service personnel from tedious routine tasks, allowing them to focus on high-value-added creative strategy work. As a result, service providers may improve productivity and expand output beyond the limits imposed by purely labor-intensive production.

Second, declining marginal costs allow some services to be delivered on a much larger scale. The core competitive advantage of modern manufacturing lies in economies of scale and increasing returns. AI enables the digitalization and standardization of professional knowledge, industry experience, and certain reasoning processes. Taking generative AI (AIGC) as an example, once a specific model is initially trained, the marginal cost difference between serving 10 clients and 100,000 clients is nearly zero. This “high fixed-cost and low marginal-cost” characteristic allows some knowledge-intensive services to achieve greater scalability similar to that observed in manufacturing industries.

Third, AI reduces geographical and time-related constraints on service delivery, thereby increasing the tradability of services. As articulated in Baldwin and Forslid’s (2020) “global robot” theory [2], the integration of AI and big data reduces the dependence of many services on physical delivery channels, allowing services to be delivered more efficiently across borders. AI-driven intelligent systems achieve precise matching of cross-border supply and demand, substantially reducing overseas market search costs and information frictions arising from cross-language communication [4]. This may improve total factor productivity in the service sector and reduce some transaction costs associated with international market participation.

In summary, AI may alleviate some of the scalability constraints traditionally associated with service industries and support the expansion of digital service production at lower marginal costs, laying the most critical material foundation for subsequent large-scale capacity spillover into international markets.

3.2 Evolution of Export Necessity Under the Trade Identity: From “domestic Demand Fulfillment” to “capacity Spillover”

In open macroeconomic models, fluctuations in net exports (NX) are not only constrained by exchange rate volatility or external trade barriers; their endogenous root lies in the structural gap between China’s total output

(Y) and total absorption ($C+I+G$). For a long time, China has exhibited significant trade deficits in many core service sectors. This essentially reflects the fact that the actual supply capacity of domestic services (Y) has lagged behind the rapidly expanding demand (C) from consumption upgrading and industrial upgrading investment (I). Under traditional industrial models, the growth curve of service capacity exhibits typical linear and modest characteristics, making it difficult to respond quickly to rapid increases in demand. As a result, China has relied more heavily on imported services in certain sectors.

The introduction of artificial intelligence technology, by changing service production processes, may accelerate the growth of service capacity and improve production efficiency. In the initial stage, AI first enhances the total factor productivity of local service enterprises through automation and intelligent tools, effectively reduces service delivery costs, and enables Chinese firms to close the gap in meeting domestic consumption and investment needs [9]. As productivity continues to improve, leveraging the marginal cost reduction effect unleashed by AI, the growth of service capacity (Y) may gradually exceed the growth of domestic absorption. Within the dynamic evolution path of the trade identity framework, when capacity (Y) surpasses the threshold of domestic absorption, net exports (NX) will inevitably exhibit a convergence trend from negative to positive. Under such conditions, excess service capacity may support the expansion of exports into international markets [10].

However, it is important to distinguish between simple overcapacity and internationally competitive production capacity. Within the trade identity framework, a simple increase in output scale (i.e., traditional overcapacity) is not sufficient to automatically convert into effective exports in international markets. According to classic heterogeneous firm trade theory [10], firms entering international markets usually face substantial fixed and sunk costs. Only enterprises with significant “technological advantage” and high productivity can remain internationally competitive and profitable after bearing these additional costs. Therefore, service exports depend not only on output scale but also on competitiveness and productivity, but rather the active spillover of advanced productive forces. In other words, exports can only be achieved when a clear technological advantage is attained; this is an important condition for successful participation in international markets. The core strategic value of artificial intelligence lies not only in expanding the scale of service output (Y), but more importantly in endowing domestic enterprises with stronger technological capabilities in international markets through advanced algorithms, data-processing capabilities, and large AI models. Only when “capacity surplus” is combined with absolute “technological innovation advantage” can it truly translate into stronger competitiveness in digital service exports.

China’s manufacturing development experience suggests that the establishment of export competitiveness generally follows the development path of “capacity expansion → domestic market saturation → export expansion.” At present, artificial intelligence is increasing the scalability and standardization of some service activities, similar to those in manufacturing. Expanding service capacity through AI aims not only to better satisfy domestic market demand, but also to follow the endogenous logic of macroeconomic identities by pushing digitally enabled service capacity—with absolute technological advantage and efficient operation—into the global trade system, thereby improving China’s position in international service trade.

3.3 Amplification Effect of Super-Large Economies: From Domestic Application Scenarios to Global Competitive Barriers

If the trade identity constitutes the theoretical cornerstone of service export expansion, then the scale of the domestic economy may help translate these theoretical advantages into international competitiveness. Because AI development depends heavily on data resources and application scenarios, this acceleration and amplification effect is primarily manifested in the following three strategic dimensions:

First, large-scale data resources and continuous model iteration may generate economies of scale. As the world’s largest single consumer market, China continuously generates massive volumes of interaction data across a wide range of service formats. These large-scale data resources provide favorable conditions for AI model training by domestic firms, improving model performance while reducing unit costs over time. This economic characteristic enables China’s domestically incubated AI systems to rapidly establish asymmetric competitive advantages.

Second, a highly complex demand ecosystem provides favorable conditions for the testing and improvement of digital services. The demand structure of China’s super-large market is exceptionally deep,

spanning from high-end knowledge-intensive services to basic digital services, thereby providing extensive early-stage application scenarios for AI technologies. In the process of meeting China's massive and rapidly evolving demands, enterprises continuously optimize the robustness of their algorithms. When these digital service products, having undergone full trial-and-error and rigorous iteration, are introduced to international markets, they inherently establish significant technological advantages and barriers to entry relative to similar services from other countries.

Third, the super-large market helps spread sunk costs across a larger market, which may help reduce aspects of the global digital divide. The construction of AI's underlying infrastructure involves extremely high fixed investments. The core advantage of a super-large economy lies in its vast domestic consumer market, which can rapidly absorb and amortize the aforementioned R&D and infrastructure costs. Once underlying general large models achieve break-even in the Chinese market, the additional marginal delivery costs of deploying their cloud capabilities overseas approach zero. This industrial economic structure of "high fixed costs, extremely low marginal costs" may strengthen the pricing power of Chinese enterprises in international markets, thereby improving the price competitiveness of service exports and increasing barriers to entry.

4. Policy Recommendations

Based on the findings of the above analysis, this paper offers the following targeted policy recommendations for government decision-making bodies and micro-level enterprises:

First, further improve digital infrastructure development. Given that network transmission speed and underlying computing power directly constrain the practical effectiveness of AI empowerment, reducing the digital divide is important for improving the inclusiveness of international digital trade. The government should systematically increase strategic investment in 5G/6G communication networks, large-scale intelligent computing centers, and other facilities, reduce enterprises' network factor usage costs, and provide infrastructure support for the development of service exports.

Second, increase the effectiveness of R&D investment. In line with the empirical findings on the moderating effect of innovation levels, micro-level enterprises should not limit themselves to the external procurement of basic application tools. Instead, they should actively pursue "intelligent empowerment" strategies, continuously increase resource allocation to core underlying algorithms, and effectively strengthen the role of innovation in improving long-term competitiveness.

Third, actively participate in the development of international digital trade rules. Facing increasingly intensified global barriers to digital service trade and regulatory restrictions on cross-border data flows [5], China should further align with high-standard international trade rules, work to eliminate unreasonable regulatory constraints, and create a freer and more secure institutional environment for cross-border data flows to support the international expansion of domestic enterprises.

5. Conclusion

Integrating macroeconomic trade statistics with micro-level empirical findings, this study demonstrates that artificial intelligence technology has a significant positive effect on the expansion of global service exports through three core mechanisms: reducing transaction costs, enhancing production efficiency, and promoting service innovation. The 2023 macroeconomic data showing that digitally deliverable service exports accounted for 54.2% of total global service exports is consistent with this trend. More importantly, artificial intelligence contributes not only to the expansion of service trade; it is also a core engine for elevating the technological complexity of service exports. It is also contributing to structural changes in international service trade—from labor-intensive activities toward more knowledge- and technology-intensive activities.

References

- [1] World Bank. (2024). Digital progress and trends report 2023. World Bank.
- [2] Baldwin, R., & Forslid, R. (2020). Globotics and development: When manufacturing is jobless and services are tradable (*NBER Working Paper No. 26731*). National Bureau of Economic Research.

- [3] Goldfarb, A., & Tucker, C. E. (2019). Digital economics. *Journal of Economic Literature*, 57(1), 3–43.
- [4] Agrawal, A., Gans, J., & Goldfarb, A. (2019). *The Economics of Artificial Intelligence: An Agenda*. University of Chicago Press.
- [5] China Academy of Information and Communications Technology (CAICT). (2024). Annual observation report on global digital economic and trade rules.
- [6] Yao, T. T., & Bao, Y. N. (2024). Does artificial intelligence promote the improvement of technological complexity of service exports? *Journal of Zhongnan University of Economics and Law*, (02), 149–160.
- [7] Tao, M. Y. (2024). Artificial intelligence, innovation and technological complexity of service exports. *Statistics and Application*, 13(06), 218–226.
- [8] Acemoglu, D., & Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6), 1488–1542.
- [9] Brynjolfsson, E., Li, D., & Raymond, L. (2023). Generative AI at work (*NBER Working Paper* No. w31161). National Bureau of Economic Research.
- [10] Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6), 1695–1725.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Analysis on the Formation Causes of AH Premium and the Reverse Premium Anomaly of Leading Companies

Yuexing Zhu*

School of Economics, Jinan University, Jinan 510000, China

**Corresponding author: Yuexing Zhu.*

Abstract

As the interconnection mechanism of Shanghai-Shenzhen-Hong Kong Stock Connect keeps improving, an increasing number of enterprises choose to go public simultaneously in the A-share and H-share markets. However, the long-standing problem of “same stock with different prices” and the persistent premium of A-shares still exist. Taking the valuation difference between A-shares and H-shares as the research object, this paper uses literature review, comparative analysis and case studies to explore the formation mechanism and fluctuation characteristics of AH share premium from the perspectives of market structure, trading system, liquidity, investor sentiment and exchange rate changes. It also focuses on the reverse premium anomaly of leading enterprises such as CATL. The results show that structural differences between the two markets are the main cause of AH share premium, and macro liquidity and capital flows significantly affect the trend of the premium. Meanwhile, industry leaders with strong fundamentals and high information transparency are more likely to show H-share reverse premium. This study can provide references for cross-market investment decisions, corporate capital operation and the institutional optimization of capital markets.

Keywords

AH premium rate, reverse premium, core leading companies

1. Introduction

With the continuous improvement of the Shanghai-Shenzhen-Hong Kong Stock Connect mechanism, A+H dual listing has become an important financing channel for domestic enterprises, and the linkage between the two capital markets has been significantly enhanced. Nevertheless, the price divergence of the same share in the two markets has existed for a long time. Generally, A-shares maintain a sustained premium over H-shares, and the premium level changes with market sentiment, policy adjustments and hot events. Some stocks even show obvious H-share premium and price inversion.

The core reason for such valuation divergence lies in the institutional and structural differences between the two markets. The Hong Kong stock market is dominated by institutional and overseas investors who prefer value investing, while the A-share market has a higher proportion of retail investors, leading to stronger volatility in market sentiment. In addition, the two markets differ greatly in trading mechanism, trading hours and price limit rules. Coupled with exchange rate fluctuations, liquidity differences and cross-market information asymmetry, the price spread between A-shares and H-shares persists and fluctuates sharply. Under the background of rising uncertainty in overseas financing of Chinese concept stocks, the strategic position of

the Hong Kong stock market has become more prominent, so it is of great practical value to conduct an in-depth study on the formation mechanism of valuation differences between A-shares and H-shares.

Most existing studies explain AH share premium from macro perspectives such as market segmentation, liquidity and investor preferences, but pay less attention to practical market factors such as mismatched trading hours and short-term capital speculation. The analysis on special phenomena such as valuation inversion and periodic premium fluctuation is also insufficient. Based on the actual trading characteristics of the two markets, this paper systematically analyzes the causes of valuation divergence through a combination of qualitative and quantitative methods, which enriches the theories of cross-market asset pricing and market segmentation, and provides a refined theoretical supplement for AH share valuation research.

For investors, the conclusions help to clarify the pricing logic of the two markets, identify the driving factors of premium fluctuations, optimize cross-market investment decisions and reduce investment risks. For enterprises, it can provide a reference for A+H dual-listed companies to choose financing timing, formulate listing strategies and optimize capital operation. For regulators, the conclusions can lay a foundation for improving the interconnection mechanism between the two markets, optimizing trading rules, weakening unreasonable valuation deviation and improving the overall pricing efficiency of the capital market, so as to boost the standardized and internationalized development of China's capital market.

A+H dual listing: An enterprise is listed on both the mainland A-share market and the Hong Kong H-share market. A-shares are traded in RMB and H-shares in Hong Kong dollars. The two types of shares have the same rights.

AH share premium and reverse premium: The price spread is usually measured by the premium rate, and the core formula is:

$$(\text{A-share price (CNH)}/\text{H-share price (HKD)}) \times \text{exchange rate} - 1) \times 100\% \quad (1)$$

A positive premium rate means that A-shares are at a premium over H-shares, which is the normal state of the market. A negative premium rate means that H-shares are at a premium over A-shares, which is the reverse premium or valuation inversion anomaly studied in this paper.

This paper employs three core methods: a literature review consolidates theoretical foundations related to market segmentation, cross-market pricing and AH share premium; a comparative analysis of trading rules, investor structures and cross-market premium differences identifies divergence drivers; and case studies of event-driven abnormal stocks concretely examine how capital and sentiment shape valuation spreads.

The innovations of this paper are reflected in research perspective and methodology. In terms of perspective, it breaks through the limitation that traditional research focuses on macro factors, and incorporates practical factors such as mismatched trading hours, exchange rate fluctuation and short-term event speculation, which is more in line with the current operation of the AH share market. In terms of methodology, it combines qualitative case analysis and quantitative data test, which can not only explain the internal logic of valuation difference, but also quantify the impact degree of core factors. It can effectively explain diverse market phenomena such as normal premium, short-term fluctuation and valuation inversion, forming an effective supplement to existing research.

2. Literature Review

Scholars at home and abroad have carried out a lot of research on the premium phenomenon of A+H dual-listed shares. This chapter systematically sorts out and reviews the existing literature from three aspects: classical explanatory theories, exchange rate factors and other influencing factors.

Focusing on the valuation difference between A-shares and H-shares, this paper first combs relevant literature, defines the research scope and summarizes the evolutionary characteristics of AH share premium. Second, it qualitatively analyzes the formation logic of valuation divergence from the dimensions of trading system, market structure and exchange rate fluctuation. Third, it compares the overall market premium and individual stock premium performance, and analyzes the causes of short-term premium fluctuations combined with hot event cases. Finally, it quantifies the impact of key variables through data statistics, summarizes the

research conclusions and puts forward relevant suggestions to improve the pricing mechanism of the two capital markets.

2.1 Classical Explanatory Theories

2.1.1 Market Segmentation Theory

In the 1970s, foreign scholars found that the share prices of the same company listed in different markets were significantly different, violating the Law of One Price in traditional economics. Solnik (1973) [1] first proposed the market segmentation theory, which has been continuously improved in follow-up studies, and gradually formed a complete theoretical system including liquidity difference hypothesis, information asymmetry hypothesis, demand elasticity difference hypothesis and risk preference difference hypothesis.

Peng (2024) [2] took 98 enterprises in the Hang Seng AH Share Premium Index as samples, confirming that A-share and Hong Kong stock market indices have strong explanatory power for AH share premium rate. The increase in daily turnover of northbound and southbound trading can weaken the effect of market segmentation and narrow the price spread between A-shares and H-shares.

2.1.2 Liquidity Difference Hypothesis

The worse the liquidity of an asset, the higher the rate of return required by investors to obtain liquidity risk compensation. Amihud & Mendelson (1986) [3] showed that investors require additional compensation for stocks with poor liquidity.

From the perspective of investor structure, the A-share market is dominated by retail investors with high trading frequency and strong liquidity, while the Hong Kong stock market is dominated by institutional investors with relatively rational and stable trading, and its liquidity is weaker than that of A-shares, thus forming AH share premium in most cases.

2.1.3 Information Asymmetry Hypothesis

Information asymmetry means that A-share and H-share investors differ in the difficulty, cost and accuracy of information acquisition, leading to price deviation of AH shares. Most A+H dual-listed companies focus their main businesses in the mainland, so A-share investors can obtain information more conveniently and at a lower cost, promoting A-share premium over H-shares.

Zheng (2025) [4] found that the AH share premium rate factor has significant differential impacts on the expected prices of stocks in different industries, and the larger the company's market capitalization, the lower the relative AH premium rate, which may be caused by information asymmetry across industries and company sizes.

2.2 Investor Sentiment

Baker & Stein (2004) [5] pointed out that investor sentiment reflects the deviation between investors' value judgment and the real value of assets. Applied to the AH share market, Qing (2023) [6] found that the investor sentiment difference index and the Shanghai-Hong Kong Stock Connect and Shenzhen-Hong Kong Stock Connect policies have positive impacts on AH premium. Liu, L. B., & Guo, L. (2022) [7] demonstrated that the divergence in investor sentiment exerts a significantly positive effect on the AH premium of dual-listed companies. Zhang & Zhang (2025) [8] found that when investors hold extrapolative price expectations, the prices of A-shares and H-shares will diverge in equilibrium.

According to the prospect theory proposed by Tversky & Kahneman (1979) [9], investors are loss-averse and more sensitive to losses than to gains of the same amount. Therefore, market sentiment has a stronger impact on stock market returns during pessimistic periods. Taking 43 AH dual-listed companies in the Shanghai-Hong Kong Stock Connect as samples, Sun (2022) [10] confirmed that investor sentiment difference can effectively explain AH share premium, and its explanatory power is stronger during periods of sharp market fluctuations.

3. Case Introduction and Selection

3.1 Criteria and Basis for Case Selection

As of April 2026, there are 183 companies listed on both A-share and H-share markets, providing sufficient samples for this study. First, this paper selects the AH share premium index from April 18, 2018 to April 30, 2026, and uses the market-value-weighted average individual stock premium rate to calculate the index premium rate, which can better reflect the overall premium trend of AH shares at the macro level.

Second, this paper screens stocks with reverse premium in the past year to fully reflect the company's latest operation. To eliminate large fluctuations of new shares, stocks listed for less than half a year are excluded. Finally, four stocks are selected: CATL, Hengrui Medicine, China Merchants Bank and WuXi AppTec. CATL is taken as the main case because of its significant reverse premium rate and representativeness.

Note: All data are as of April 30, 2026.

3.2 Data Source and Case Background

All data are from www.10jqka.com.cn.

As shown in Figure 1, From 2018 to 2026, the overall AH share premium shows a trend of rising, fluctuating at a high level and then falling. From 2018 to 2020, it was affected by the COVID-19 pandemic. From 2020 to 2022, liquidity in the A-share market was loose, while the H-share market was under pressure from overseas liquidity tightening. Coupled with divergent market sentiment under the pandemic, the AH premium soared from about 20% to around 60%. It fluctuated at a high level from 2022 to 2024, and narrowed rapidly from 2024 to 2026 due to Hong Kong stock valuation recovery and the inflow of southbound capital.

Figure 1: AH share premium index, 2018.4.18–2026.4.30.



4. Case Analysis: Causes of the AH Share Premium Anomaly

4.1 Common Characteristics of Case Companies

From the above data, stocks with reverse premium all have a market capitalization of more than 30 billion CNY, and they are industry giants with economic moats and stable profitability. Mature business models largely solve the problem of information asymmetry. Under the overall low liquidity of the Hong Kong stock market, foreign investors are more willing to invest in these companies, thus forming reverse premium.

4.2 Causes of the Anomaly Based on Theoretical Framework

From the perspective of market segmentation theory, the A-share and Hong Kong stock markets have long been clearly segmented in investor structure and trading system, and most companies show A-share premium. The four enterprises selected in this paper are all industry leaders with stable asset quality and high global recognition, which effectively weaken the valuation gap caused by market segmentation. The Hong Kong stock market, dominated by overseas institutional investors, tends to follow international valuation logic, while

the A-share market is more susceptible to short-term sentiment and policies with relatively low valuation, eventually forming H-share reverse premium.

From the perspective of liquidity difference hypothesis, ordinary AH share companies usually have better A-share liquidity, thus presenting A-share liquidity premium. However, H-shares of leading enterprises have smaller tradable shares and stable institutional holdings, resulting in relatively scarce tradable chips and forming a scarcity premium instead. A-shares have larger tradable shares with no obvious liquidity advantage, reversing the pricing logic between the two markets and promoting the emergence of reverse premium.

From the perspective of information asymmetry theory, leading enterprises have standardized governance, transparent information disclosure and certain international layout, enabling overseas investors to fully understand their operation, significantly reducing cross-market information asymmetry. Foreign investors do not have to offer discounts due to information disadvantages and are more willing to give reasonable valuations, supporting the rise of H-share prices.

Combined with investor sentiment factors, A-share investors pay more attention to short-term performance and market sentiment fluctuations, making valuations vulnerable to disturbances; Hong Kong stock institutions focus more on long-term value and fundamentals with more rational pricing. The divergence in investment preferences between the two markets coupled with cross-border capital flows further narrows and eventually reverses the traditional AH premium, leading to H-share reverse premium for high-quality leaders.

4.3 Uniqueness and Comparative Analysis of Each Case

4.3.1 CATL: AI-Driven Energy Storage Expansion and Global Leadership

CATL has successfully built energy storage as the second growth curve, with substantial revenue growth in the first quarter of 2026. The growth rate of energy storage battery shipments far exceeds that of power batteries, formally forming a dual-engine development pattern of power batteries and energy storage.

The surge in AI computing demand drives the upgrading of power infrastructure, and energy storage has shifted from traditional auxiliary supporting positioning to the core energy base supporting the operation of AI data centers, greatly broadening the industry development space.

Meanwhile, relying on a mature cost transmission mechanism and large-scale advantages, the company maintains a stable gross profit margin amid raw material price fluctuations with strong profitability certainty. In addition, CATL ranks top in the global, mainland Chinese and European power battery markets, as well as the global energy storage battery market, with prominent global competitive advantages and a solid industry moat.

4.3.2 Hengrui Medicine: R&D-Driven Barriers in the Pharmaceutical Industry

It has built a closed-loop moat covering the whole innovative drug chain with high R&D investment, rich innovative drug pipeline, full technical platforms, strong commercial channels and sound finance.

4.3.3 China Merchants Bank: Deep Competitive Barriers in Retail Finance

It has formed a collaborative moat with high-quality retail customers, low-cost liabilities, wealth management, financial technology and robust risk control, featuring deep and irreplicable barriers.

4.3.4 WuXi AppTec: Leadership in Integrated CRDMO Services

The firm has established the strongest competitive position in the CXO industry via end-to-end CRDMO integration, high customer switching costs, technological and scale advantages, and strategic partnerships with global leading clients.

As shown in Table 1, CATL shows the most obvious reverse premium, which is not only due to its solid fundamentals and stable profitability from the power battery business, but also because overseas investors are optimistic about its unique energy storage technology in the era of AI infrastructure. Foreign investors reprice CATL in line with international standards based on the future AI mainstream narrative. The premium of China Merchants Bank mainly comes from foreign investors' optimism about its business model and growth potential. Its unique operating model ensures steady profit growth, and as a banking stock, it has a high dividend payout compared with the other three stocks. Thus, foreign investors revalue China Merchants Bank as a high-quality

Chinese banking stock. Hengrui Medicine has R&D barriers and sustained high R&D investment, and foreign investors are more optimistic about its internationalization potential of innovative drugs. WuXi AppTec exhibits relatively lower business growth, hence a milder reverse premium.

Table 1: Typical reverse leading companies.

Stock Name	Average Premium Rate in the Past Year (H/A)	A-share Market Cap (CNY)	H-share Market Cap (HKD)
CATL	26.92%	1.99 trillion	2.77 trillion
China Merchants Bank	7.16%	965.16 billion	1.19 trillion
Hengrui Medicine	6.99%	357.61 billion	431.42 billion
WuXi AppTec	1.55%	326.87 billion	405.79 billion

5. Conclusion

This paper studies the AH premium and reverse premium phenomena, finding that differences in investor structure, trading system, liquidity and market sentiment between the A-share and Hong Kong stock markets are the core causes of the long-term AH share premium, while exchange rates and capital flows will aggravate premium fluctuations. Industry leaders, favored by foreign investors in the Hong Kong stock market due to complete information and stable profitability, present reverse premium, indicating that high-quality fundamentals can weaken the impact of market segmentation. This study has limitations such as limited sample coverage and dominance of qualitative analysis, failing to accurately characterize the impact degree of each factor.

For investors, it is necessary to participate in cross-market investment rationally by combining the differences between the two markets and corporate fundamentals. For regulators, efforts should be made to continuously improve the interconnection mechanism, promote the convergence of trading rules and information disclosure standards, and enhance cross-market pricing efficiency.

Future research can expand the sample scope, introduce quantitative models to test premium influencing factors, and combine high-frequency data and sentiment surveys to deeply analyze the impact of capital behavior and expectations on price spreads, so as to further deepen cross-market asset pricing research.

References

- [1] Solnik B H. (1977). Testing international asset pricing: Some pessimistic views. *The Journal of Finance*, 32(2), 503–512.
- [2] Peng Z Q. (2024). Research on Influencing Factors of AH Share Premium.
- [3] Amihud Y, Mendelson H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17(2), 223–249.
- [4] Zheng X. AH premium rate and A-share valuations: DCF analysis of cross-market predictability in dual-listed equities. 2025, 9–10.
- [5] Baker M, Stein J C. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets*, 7(3), 271–299.
- [6] Qing C H. (2023). Research on the Impact of Investor Sentiment on AH Share Premium.
- [7] Liu L B, Guo L. (2022). Can Sentiment Differences Explain AH Share Premium? 42(1), 66–80.
- [8] Zhang R, Zhang T. The AH premium: A tale of “siamese twin” stocks. 2025, 16–17.
- [9] Tversky A, Kahneman D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2), 263–291.
- [10] Sun Y. (2022). Investor Sentiment, Shanghai-Hong Kong Stock Connect and AH Share Premium.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open - access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

The Impact of Talent Introduction Intensity on Corporate Artificial Intelligence Levels: Empirical Evidence from Chinese A-Share Listed Companies

Shanlin Bi*

School of Health Management, Shandong University of Traditional Chinese Medicine, Jinan 250300, China

**Corresponding author: Shanlin Bi.*

Abstract

Talent introduction has become an important factor shaping enterprise development in recent years. Using panel data for Shanghai and Shenzhen A-share listed companies, this paper examines how talent introduction affects firms' artificial intelligence (AI) development and explores the channels through which this effect operates. The empirical findings suggest that talent introduction is associated with higher levels of AI development among firms. This relationship remains robust after a series of robustness tests and instrumental variable estimations. Further analysis indicates that talent introduction may contribute to AI development by easing financing constraints and improving workforce quality. The effects are not uniform across firms, however. They appear to be more pronounced depending on characteristics such as pollution status, regional location, and industry affiliation, particularly in the manufacturing sector. Overall, the findings highlight the role of talent introduction in supporting corporate AI development and provide evidence that may be useful for the design of talent-related policies and the promotion of the AI industry.

Keywords

talent introduction intensity, corporate artificial intelligence level, financing constraints, employee quality, propensity score matching

1. Introduction

The Report to the 20th National Congress of the Communist Party of China emphasized the importance of education, scientific and technological self-reliance, and talent-driven development. It also highlighted the need to strengthen the national strategic talent workforce, providing policy support for talent development in key sectors such as artificial intelligence (AI). In recent years, a new wave of technological change and industrial transformation has accelerated the development and diffusion of emerging technologies [1,2]. Among them, AI has attracted considerable attention because of its growing role in industrial upgrading and business innovation. As AI technologies become increasingly integrated into production and management activities, firms' ability to develop and apply AI is often regarded as an important source of competitiveness. In this process, talent plays a particularly important role. The quality, structure, and allocation of human resources may affect how effectively firms adopt and utilize AI-related technologies.

Despite the rapid expansion of China's AI industry, shortages of qualified talent remain a major challenge. Competition for highly skilled AI professionals, especially those engaged in algorithm development and large-model research, has intensified in recent years, leading to substantial salary premiums. Existing estimates suggest that China's AI talent gap may reach 4 million by 2030, with shortages being especially severe in basic research and interdisciplinary fields [3]. At the same time, mismatches between university training and industry demand, rising recruitment costs, and difficulties in integrating newly recruited personnel have limited the efficiency of talent allocation and constrained the development of AI capabilities within firms [4,5].

From the perspective of enterprise practice, the relationship between talent introduction and AI development is not always straightforward. Recruiting highly qualified personnel does not necessarily lead to immediate improvements in technological capability. In many cases, firms still face challenges in transforming talent advantages into measurable innovation outcomes. The effectiveness of talent introduction may also differ across firms because of variations in industrial characteristics, regional environments, and organizational conditions. Some firms continue to rely heavily on high-salary recruitment strategies, while mechanisms for integrating talent into long-term AI development remain relatively underdeveloped. As a result, talent investment and actual AI performance do not always move in parallel.

At the global level, the movement of high-end AI talent from universities and research institutions toward industry has accelerated, and competition for skilled professionals has become increasingly intense. Under these circumstances, talent allocation efficiency is emerging as an important factor influencing the diffusion and application of AI technologies [6]. Against this backdrop, examining the impact of talent introduction on corporate AI development from a micro-level perspective is of both theoretical and practical significance. Such an analysis can contribute to the literature on technological innovation and talent allocation, while also providing evidence for firms seeking to improve talent strategies and for policymakers designing talent-related policies.

The measurement of corporate AI development has received increasing attention in recent years, and a number of approaches have been proposed in the literature. Existing studies often rely on patent information, textual analysis, or technology-related investment indicators to capture firms' AI activities. For instance, An Tongliang et al. (2025) [7] measured corporate AI technology levels using hierarchical patent identification and network embedding methods. Other studies have used AI-related keywords disclosed in annual reports to construct indicators of AI application intensity at the firm level. These approaches have provided useful tools for examining the adoption and development of AI technologies within enterprises.

Building on these measurement methods, recent research has explored a range of factors associated with corporate AI development. Existing evidence suggests that digital infrastructure, technological innovation, and R&D investment can facilitate firms' intelligent transformation and AI-related activities [8,9]. Nevertheless, relatively little attention has been paid to the role of talent introduction, despite the fact that talent is widely regarded as an important resource for technological development and innovation. Research concerning AI-related talent has mainly focused on the broader effects of talent policies on industrial development and labor market outcomes. Comparatively fewer studies have examined whether talent introduction influences firms' AI development and, if so, through what channels this influence operates. As a result, evidence from the micro-level perspective of enterprises remains limited.

To address this gap, this paper investigates the relationship between talent introduction and corporate AI development using panel data from Chinese A-share listed companies. Particular attention is paid to the mechanisms through which talent introduction may affect firms' AI development, and empirical tests are conducted to examine these potential channels. This study contributes to the existing literature in several respects. First, it examines the mechanisms linking talent introduction to corporate AI development, focusing on the roles of financing constraints and employee quality. Second, unlike much of the existing literature that emphasizes policy effects at the industry or regional level, this paper adopts a firm-level perspective and considers both the quantity and quality dimensions of introduced talent. Third, by using micro-level panel data and conducting a series of robustness and endogeneity tests, the study provides additional evidence on the relationship between talent introduction and firms' AI development.

2. Theoretical Analysis and Research Hypotheses

2.1 Talent Introduction Intensity and Corporate Artificial Intelligence Levels

In the context of the digital economy, artificial intelligence (AI) has become increasingly important for firms seeking to improve productivity, enhance innovation capacity, and maintain competitive advantages. As the development and application of AI technologies depend heavily on human capital, talent introduction may play an important role in shaping firms' AI development. Previous studies suggest that relying solely on employees in explicitly AI-related positions is often insufficient. Firms may also benefit from attracting individuals with complementary skills and implicit AI-related capabilities, which can broaden the scope of AI adoption and facilitate the integration of AI technologies into business operations [10].

One possible channel is innovation enhancement. The introduction of highly skilled personnel can provide firms with access to new knowledge, technical expertise, and industry experience. In the field of AI, talents with backgrounds in algorithm development, machine learning, or industry-specific applications may help firms overcome technological constraints and accelerate the adoption of new technologies. In addition, newly introduced employees can facilitate knowledge diffusion within organizations, encourage collaboration across departments, and support the transformation of technological knowledge into practical applications. These effects may contribute to improvements in firms' AI development over time.

At the same time, the benefits of talent introduction are not automatic. Whether newly recruited employees can effectively contribute to AI development depends on firms' organizational capacity and resource allocation. High recruitment costs, difficulties in integrating new employees into existing teams, and mismatches between talent structure and business needs may reduce the effectiveness of talent introduction [11]. Excessive reliance on external recruitment may also weaken incentives for internal talent cultivation and limit the accumulation of firm-specific knowledge [12]. These factors suggest that the impact of talent introduction may vary across firms and may depend on the extent to which enterprises are able to effectively absorb and utilize newly acquired human capital.

Overall, although talent introduction may face certain constraints in practice, it is generally expected to provide firms with valuable knowledge resources and innovation capabilities that support the development and application of AI technologies. Accordingly, this paper proposes the following hypothesis:

Hypothesis 1: Talent introduction is positively associated with corporate AI development.

2.2 Talent Introduction Intensity, Financing Constraints, and Corporate Artificial Intelligence Levels

Financing constraints may represent an important channel through which talent introduction influences corporate AI development. By improving firms' access to financial resources, talent introduction can create more favorable conditions for AI-related investment and technological upgrading.

Talent introduction may help reduce financing constraints in several ways. The recruitment of highly educated and highly skilled employees, particularly those with expertise in frontier technologies, can signal stronger growth potential and innovation capacity to external investors. Such signals may improve firms' reputation in capital markets and increase the availability of external financing. In addition, high-quality talent can contribute to improvements in corporate governance, operational efficiency, and information disclosure, thereby reducing information asymmetry and lowering financing costs. These factors may expand financing channels and provide greater financial support for AI-related R&D activities and technology investment [13].

At the same time, adequate financial resources are often required for the development and application of AI technologies. AI projects typically involve substantial expenditures on infrastructure, software and hardware investment, technological research, and scenario-based implementation. When financing constraints are eased, firms may be better positioned to sustain long-term and high-risk innovation activities and to allocate resources to AI-related projects. Greater financial flexibility can also facilitate the integration of AI technologies into production, management, and decision-making processes, thereby supporting the further development of corporate AI capabilities [14].

Based on this reasoning, the following hypothesis is proposed:

Hypothesis 2: Talent introduction promotes corporate AI development by alleviating financing constraints.

2.3 Talent Introduction Intensity, Employee Quality, and Corporate Artificial Intelligence Levels

Talent introduction may influence corporate AI development not only through the direct recruitment of technical personnel but also through improvements in overall employee quality. As AI technologies become increasingly integrated into business operations, the quality of the workforce may affect firms' ability to adopt, apply, and further develop these technologies.

One way in which talent introduction affects employee quality is through the optimization of workforce structure. By recruiting highly educated and skilled employees, firms can supplement existing human capital and address capability gaps in areas such as artificial intelligence, data analysis, and digital technologies. In addition, newly recruited talent may bring new knowledge, technical experience, and working methods that can gradually spread within the organization. Through day-to-day collaboration and learning, these effects may contribute to broader improvements in employee skills and professional capabilities [15].

At the same time, employee quality is closely related to the implementation and use of AI technologies. Employees with stronger technical backgrounds and digital skills are generally better able to understand and utilize AI-related tools in their daily work. A more skilled workforce may also facilitate the integration of AI technologies into different business processes and support the exploration of new application scenarios. In addition, employees with higher levels of expertise are often more capable of identifying practical problems and adapting technological solutions to business needs, which may improve the effectiveness of AI applications and innovation activities within firms [16].

Taken together, talent introduction may contribute to corporate AI development by improving the quality of the workforce and strengthening firms' capacity to absorb and apply AI-related knowledge. Accordingly, the following hypothesis is proposed:

Hypothesis 3: Talent introduction promotes corporate AI development by improving employee quality.

3. Research Design

3.1 Sample Selection and Data Sources

The sample consists of Chinese A-share listed companies during the period 2015–2024. Data on firms' digitalization and intelligent transformation are obtained from the MarkData database, while firm characteristics, financial information, and corporate governance variables are collected from the CSMAR database. Following common practice, ST and *ST firms, financial enterprises, and observations with missing values for key variables are excluded. The final sample contains 41,098 firm-year observations. To reduce the influence of extreme values, all continuous variables are winsorized at the 1% level.

3.2 Variable Definitions

(1) Independent Variable

Talent introduction intensity (Intensity). Following Chen Xi et al. [17–21], talent introduction intensity is measured using the frequency of talent-related recruitment information extracted from textual data. The variable is constructed by adding one to the original value and then taking its natural logarithm.

(2) Dependent Variable

Corporate AI development (AI). Following Sun Wenyan et al. [22], corporate AI development is measured based on the degree of AI adoption at the firm level.

(3) Mediating Variables

Financing constraints (SA). Financing constraints are measured using the SA index proposed in the literature [23,24]. The index is calculated based on firm size and firm age and is widely used as a proxy for financing constraints. A lower absolute value of the SA index indicates weaker financing constraints. In the

empirical analysis, the SA index is used to examine whether financing constraints serve as a channel through which talent introduction affects corporate AI development.

Employee quality (Quality). Employee quality is measured by the proportion of employees holding a master's degree, calculated as the number of employees with master's degrees divided by total employment. Compared with broader educational indicators, this measure reflects the share of employees with relatively high levels of professional training and technical expertise [25]. It is used to capture the role of workforce quality in the relationship between talent introduction and corporate AI development.

(4) Control Variables

Following the existing literature, several firm-level characteristics are included as control variables, including listing age (ListAge), the Herfindahl–Hirschman Index (HHI), firm size (Size), quick ratio (Quick), return on assets (ROA), revenue growth rate (Growth), book-to-market ratio (BM), ownership concentration (Top1), and board size (Board). Specifically, ListAge is measured as the number of years since listing. HHI is calculated as the sum of squared market shares of firms within the same industry and is used to capture industry concentration. Size is measured by the natural logarithm of total assets. Quick is defined as the ratio of quick assets to current liabilities. ROA is calculated as net profit divided by average total assets. Growth is measured by the annual growth rate of operating revenue. BM is the ratio of book value to market value. Top1 represents the shareholding ratio of the largest shareholder, and Board is measured by the number of directors on the board.

3.3 Regression Model Specification

To examine the relationship between talent introduction and corporate AI development, the following econometric models are constructed:

$$AI_{i,t} = \alpha_0 + \alpha_1 Intensity_{i,t} + \alpha_2 controls_{i,t} + \delta_{year} + \eta_{Firm} + \varepsilon_{i,t} \quad (1)$$

$$SA_{i,t} = \alpha_0 + \alpha_1 Intensity_{i,t} + \alpha_2 controls_{i,t} + \delta_{year} + \eta_{Firm} + \varepsilon_{i,t} \quad (2)$$

$$employee_{i,t} = \alpha_0 + \alpha_1 TechConv_{i,t} + \alpha_2 controls_{i,t} + \delta_{year} + \eta_{Firm} + \varepsilon_{i,t} \quad (3)$$

where i and t denote firm and year, respectively; AI represents the corporate artificial intelligence level; $Intensity$ denotes talent introduction intensity; SA is the financing constraints index; $employee$ represents employee quality level; $controls$ denotes the set of control variables; δ_{year} represents year fixed effects; η_{Firm} represents firm fixed effects; and ε is the random disturbance term.

4. Empirical Analysis

4.1 Descriptive Statistics

The descriptive statistics are reported in Table 1.

Table 1: Descriptive Statistics

Variable	N	Min	Max	Mean	p50	SD
	41098	8.038	15.740	12.600	12.650	1.285
Intensity	41098	0.788	5.013	3.364	3.651	1.092
ListAge	41098	0.693	3.367	2.186	2.303	0.788
HHI	41098	0.013	0.582	0.088	0.062	0.094
Size	41098	19.840	26.300	22.250	22.060	1.304
Quick	41098	0.187	14.120	1.955	1.238	2.220
ROA	41098	-0.257	0.194	0.033	0.034	0.064
Growth	41098	-0.585	2.401	0.155	0.095	0.397
	41098	0.114	1.199	0.621	0.618	0.253
Top1	41098	0.082	0.739	0.336	0.312	0.148
Board	41098	1.609	2.639	2.117	2.197	0.198

The dependent variable, corporate AI development (AI), includes 41,098 firm-year observations. The minimum value is 8.038, the maximum is 15.740, the mean is 12.600, and the standard deviation is 1.285. Overall, there is noticeable variation in AI development across firms.

The key explanatory variable, talent introduction intensity (Intensity), also consists of 41,098 observations. Its minimum value is 0.788, maximum is 5.013, mean is 3.364, and standard deviation is 1.092. The results suggest that firms differ in the extent of talent introduction, with both relatively high and relatively low values observed in the sample.

4.2 Benchmark Regression Results

The benchmark regression results are reported in Table 2. The table presents the estimated effect of talent introduction intensity on corporate AI development, controlling for year fixed effects, province fixed effects, and a set of firm-level covariates. The results indicate that the coefficient on talent introduction intensity is 0.030 and is statistically significant at the 1% level, suggesting a positive association between talent introduction and corporate AI development. In terms of economic magnitude, the estimate implies that a one-unit increase in talent introduction intensity is associated with an increase of about 0.030 units in corporate AI development, holding other variables constant. This result suggests that talent introduction is positively related to firms' intelligent transformation. Overall, the finding is consistent with Hypothesis 1.

Table 2: Benchmark Regression Results

VARIABLES	(1)
	AI
Intensity	0.030***
	(2.765)
ListAge	0.225***
	(10.517)
HHI	0.178**
	(2.100)
Size	0.077***
	(6.774)
Quick	-0.022***
	(-6.445)
ROA	-0.393***
	(-4.461)
Growth	-0.061***
	(-5.371)
BM	0.179***
	(5.166)
Top1	0.104
	(1.459)
Board	0.002
	(0.046)
Constant	10.145***
	(41.812)
Observations	41,098
R-squared	0.666
YearFE	YES
ProvinceFE	YES
Adj.R ²	0.618

Note: ***, *, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively. T-statistics (or Z-statistics) are reported in parentheses. The same applies below.

4.3 Robustness Tests

The robustness test results are reported in Table 3. A series of alternative specifications are used to examine the stability of the baseline results, including adding additional control variables, propensity score matching

(PSM), excluding observations from special periods, excluding municipalities, and applying an alternative winsorization method.

After including additional firm-level controls (Column 1), the coefficient of Intensity is 0.0230 and remains statistically significant at the 5% level, consistent with the baseline results. In Column (2), the PSM approach with 1:2 nearest-neighbor matching is applied. The balance test shows that standardized differences for all covariates are below 5% after matching. The estimated coefficient of Intensity is 0.0240 and remains significant at the 5% level. In Column (3), after excluding observations from special periods, the coefficient increases to 0.0368 and is significant at the 1% level. In Column (4), when municipality samples are excluded, the coefficient remains positive but is no longer statistically significant. In Column (5), using an alternative winsorization method, the coefficient is 0.0330 and remains significant at the 1% level.

Overall, the results across different specifications are broadly consistent with the baseline findings, suggesting that the relationship between talent introduction intensity and corporate AI development is relatively stable.

Table 3: Robustness Test Regression Results

	(1)	(2)	(3)	(4)	(5)
	Add Controls	PSM	Exclude Special Periods	Exclude Municipalities	Change Winsorization
VARIABLES	AI	AI	AI	AI	AI
Intensity	0.023** (1.999)	0.024** (2.019)	0.037*** (3.175)	0.009 (0.661)	0.033*** (3.864)
ListAge	0.232*** (9.966)	0.200*** (8.979)	0.216*** (9.122)	0.154*** (7.055)	0.167*** (11.304)
HHI	0.084 (0.943)	0.070 (0.743)	0.294*** (3.507)	0.227*** (2.738)	0.222** (2.549)
Size	0.078*** (6.586)	0.082*** (6.894)	0.124*** (10.135)	0.099*** (8.949)	0.109*** (13.137)
Quick	-0.023*** (-6.644)	-0.021*** (-5.857)	-0.024*** (-6.132)	-0.027*** (-6.940)	-0.035*** (-8.623)
ROA	-0.434*** (-4.694)	-0.452*** (-4.757)	-0.624*** (-5.901)	-0.516*** (-5.512)	-0.822*** (-8.421)
Growth	-0.063*** (-5.311)	-0.052*** (-4.349)	-0.052*** (-4.143)	-0.060*** (-4.893)	-0.058*** (-4.351)
BM	0.170*** (4.728)	0.204*** (5.640)	0.185*** (4.589)	0.193*** (5.417)	0.119*** (4.497)
Top1	0.080 (1.055)	0.077 (1.024)	0.037 (0.484)	-0.104 (-1.495)	-0.049 (-0.919)
Board	0.034 (0.806)	0.026 (0.610)	0.012 (0.273)	-0.036 (-0.882)	0.009 (0.284)
Constant	10.102*** (39.603)	10.015*** (39.300)	9.131*** (36.161)	9.930*** (42.869)	9.707*** (55.312)
Observations	38,478	36,510	34,567	36,994	46,121
R-squared	0.664	0.680	0.632	0.630	0.695
Adj.R ²	0.628	0.628	0.569	0.583	0.657

4.4 Mechanism Tests

Table 4 reports the mechanism test results on the relationship between talent introduction intensity and corporate AI development, focusing on two channels: financing constraints (SA index) and employee quality.

Column (1) presents the results with the SA index as the dependent variable. The coefficient on talent introduction intensity (Intensity) is -0.003 and is statistically significant at the 1% level. Given that the SA index is a reverse measure of financing constraints, this result indicates that higher talent introduction intensity is associated with weaker financing constraints. This pattern suggests that firms with stronger talent introduction may face fewer financing frictions, which could be related to improved access to external financing and better internal governance conditions.

Column (2) reports the results using employee quality as the dependent variable. The coefficient on Intensity is 0.069 and is significant at the 10% level. This implies a positive association between talent introduction intensity and employee quality. One possible explanation is that talent introduction increases the share of highly educated and skilled employees and may also contribute to broader improvements in workforce capabilities through knowledge diffusion within the firm.

Overall, the results are consistent with the existence of two potential channels linking talent introduction to corporate AI development, namely financing conditions and employee quality. These findings provide supporting evidence for Hypotheses 2 and 3.

Table 4: Mechanism Regression Results

VARIABLES	(1) SA	(2) employee
Intensity	-0.003*** (-3.710)	0.069* (1.814)
ListAge	-0.066*** (-43.929)	-0.644*** (-8.658)
HHI	-0.012** (-1.979)	0.821*** (2.689)
Size	0.010*** (12.409)	0.523*** (12.784)
Quick	0.003*** (12.329)	0.118*** (9.741)
ROA	-0.023*** (-3.662)	-0.911*** (-2.944)
Growth	-0.009*** (-11.070)	-0.050 (-1.263)
BM	-0.032*** (-13.174)	-0.611*** (-5.084)
Top1	0.020*** (4.013)	0.930*** (3.587)
Board	-0.012*** (-4.281)	-0.033 (-0.240)
Constant	-3.628*** (-213.225)	-6.908*** (-7.850)
Observations	41,098	29,944
R-squared	0.966	0.899
Adj.R ²	0.883	0.883

4.5 Heterogeneity Analysis

4.5.1 Heterogeneity Test Based on Enterprise Pollution Attributes

To examine whether the impact of talent introduction intensity on corporate AI development varies by pollution characteristics, the sample is divided into polluting and non-polluting firms based on industry pollution intensity, and regressions are conducted separately.

The results show that the coefficient of talent introduction intensity is 0.074 and statistically significant at the 1% level in polluting firms. In non-polluting firms, the coefficient is 0.010 and not statistically significant. The difference may be related to variation in environmental regulation pressure and transformation incentives across the two groups. Polluting firms generally face stronger regulatory constraints and more urgent emission reduction needs, which may create clearer demand for intelligent transformation and more identifiable application scenarios for AI technologies. In contrast, non-polluting firms face relatively weaker external pressure, and the role of talent introduction in driving AI development appears more limited.

4.5.2 Heterogeneity Test Based on Enterprise Registration Region

To examine whether regional differences affect the relationship between talent introduction and corporate AI development, the sample is divided into eastern and western regions based on firms' registration locations, and regressions are conducted separately.

The results show that the coefficient of talent introduction intensity is 0.037 and significant at the 5% level in the western region. In the eastern region, the coefficient is 0.013 and not statistically significant. The difference may be related to variation in regional development conditions. The eastern region generally has more developed digital infrastructure and more intense competition for talent, where firms may rely more on external cooperation and existing technological resources for AI development, reducing the marginal role of talent introduction. In contrast, firms in the western region tend to face relatively weaker talent availability and larger technological gaps, so talent introduction may play a more noticeable role in supporting AI-related development, especially when combined with local policy support.

4.5.3 Heterogeneity Test Based on Enterprise Industry Type

To examine whether industry characteristics affect the relationship between talent introduction and corporate AI development, the sample is divided into manufacturing and non-manufacturing firms based on the National Economic Industry Classification Standard, and regressions are conducted separately.

The results show that the coefficient of talent introduction intensity is 0.051 and significant at the 1% level in the manufacturing group. In the non-manufacturing group, the coefficient is -0.004 and not statistically significant. The difference may be related to variations in production and application contexts across industries. Manufacturing firms typically have more standardized production processes and more clearly defined AI application scenarios, such as intelligent manufacturing and process optimization, where talent introduction may be more directly reflected in technological upgrading. In contrast, non-manufacturing firms, especially service-oriented ones, tend to rely more on heterogeneous and scenario-based service activities, where AI applications are less standardized and the role of talent introduction appears less pronounced.

Table 5: Heterogeneity Tests

	(1)	(2)	(1)	(2)	(1)	(2)
	Polluting	Non-polluting	Eastern	Western	Manufacturing	Non-manufacturing
Intensity	0.074*** (4.344)	0.010 (1.002)	0.013 (1.169)	0.037** (2.385)	0.051*** (5.100)	-0.004 (-0.284)
ListAge	0.201*** (6.330)	0.157*** (9.401)	0.240*** (13.715)	0.023 (0.805)	0.094*** (5.528)	0.212*** (7.542)
HHI	0.687*** (2.686)	0.113 (1.158)	0.351*** (3.344)	0.036 (0.227)	0.403*** (2.998)	0.357** (2.540)
Size	0.142*** (8.038)	0.108*** (11.253)	0.104*** (10.218)	0.096*** (6.638)	0.128*** (12.995)	0.066*** (4.197)
Quick	-0.072*** (-8.301)	-0.024*** (-5.346)	-0.025*** (-5.333)	-0.064*** (-8.265)	-0.047*** (-10.297)	-0.026*** (-3.338)
ROA	-0.106 (-0.550)	-1.219*** (-10.826)	-0.997*** (-8.492)	-0.435** (-2.492)	-0.951*** (-8.580)	-0.999*** (-5.490)
Growth	0.024 (0.835)	-0.080*** (-5.390)	-0.054*** (-3.369)	-0.062*** (-2.636)	-0.066*** (-4.171)	-0.053** (-2.350)
BM	0.229*** (4.221)	0.023 (0.774)	0.088*** (2.795)	0.170*** (3.546)	0.136*** (4.561)	0.154*** (3.019)
Top1	-0.131 (-1.231)	-0.007 (-0.117)	0.152** (2.312)	-0.441*** (-4.914)	-0.080 (-1.297)	0.049 (0.460)
Board	-0.057 (-0.830)	0.061 (1.642)	0.123*** (3.057)	-0.233*** (-4.044)	-0.064* (-1.702)	0.099 (1.596)
Constant	9.096*** (24.327)	9.728*** (47.788)	9.540*** (43.703)	10.775*** (35.312)	9.395*** (45.294)	10.439*** (30.215)
Observations	10,635	35,486	32,322	13,799	29,870	16,251
R-squared	0.674	0.688	0.706	0.682	0.685	0.748
YearFE	YES	YES	YES	YES	YES	YES
ProvinceFE	YES	YES	YES	YES	YES	YES
Adj.R ²	0.645	0.645	0.645	0.645	0.715	0.715

5. Conclusions and Implications

This paper uses Chinese A-share listed companies from 2015 to 2024 to examine the relationship between talent introduction intensity and corporate AI development, as well as the related mechanisms and heterogeneity patterns. The empirical results suggest that talent introduction is generally associated with higher levels of corporate AI development. This finding remains stable across different robustness checks and after addressing potential endogeneity issues.

The mechanism analysis shows that talent introduction may work through two channels, including easing financing constraints and improving employee quality. These channels do not operate in isolation but appear to jointly support firms' ability to adopt and apply AI-related technologies. The heterogeneity results further indicate that the effect is more evident in polluting firms, firms located in western regions, and manufacturing industries. In contrast, the relationship is weaker or not statistically significant in non-polluting firms, eastern regions, and non-manufacturing sectors.

Based on these findings, several policy and managerial implications can be considered. First, firms may need to think more carefully about how talent is introduced and used. Simply increasing recruitment may not be sufficient. More attention could be given to matching talent with existing organizational needs, especially for employees with cross-disciplinary backgrounds, while also strengthening internal training and knowledge absorption. Second, policy support may still matter, but it does not need to be uniform. Regions and industries with weaker talent bases, such as western areas and manufacturing-related sectors, may require more targeted support in terms of funding and talent policies, while also improving the connection between university training and firm demand. Third, from an industry perspective, reducing unnecessary talent competition and improving channels for talent and knowledge sharing may help. At the same time, the way AI is applied still differs across industries, so it may be more practical to encourage firms to develop AI applications according to their own business contexts rather than following a single model.

References

- [1] Brynjolfsson, E., & McAfee, A. (2014). The second machine age: Work, progress, and prosperity in a time of brilliant technologies. *NZ Business*, 14(11), 1895-1896.
- [2] Jia, G. L. (2016). The third industrial revolution and industrial intelligence. *Social Sciences in China*, (06), 87-106+206.
- [3] Yu, F. X. (2026). Research on the challenges and paths of artificial intelligence talent cultivation in China. *China Development Observation*, (01), 33-38.
- [4] Hu, Y., Su, S. L., & Jiang, Q. X. (2025). Research on the cultivation path of artificial intelligence talents in China's "Double First-Class" universities from the perspective of interdisciplinary integration. *Journal of Innovation and Entrepreneurship Education*, 16(03), 138-143.
- [5] Tang, G. Y., & Sun, Q. (2024). Crossing technical boundaries: The dilemma and coping strategies of AI recruitment. *Quarterly Journal of Management*, 9(03), 37-46+143.
- [6] Akcigit, U., Chikis, C. A., Dinlersoz, E., & Goldschlag, N. (2026). *Attention (and money) is all you need: Why universities are struggling to keep AI talent* (NBER Working Paper No. 34964). National Bureau of Economic Research.
- [7] An, T. L., & Wu, Z. Y. (2025). Hierarchical measurement of corporate artificial intelligence technology level and its innovation effect: From the perspective of global knowledge network. *China Industrial Economy*, (06), 81-100. <https://doi.org/10.19581/j.cnki.ciejjournal.2025.06.005>.
- [8] Li, X. Y., & Ye, C. S. (2026). New measurement methods of artificial intelligence and their applications. *The Journal of World Economy*, 49(04), 68-110. <https://doi.org/10.19985/j.cnki.cassjwe.2026.04.009>.
- [9] Yu, M. G., He, M. M., & Zhang, M. M. (2024). Talent introduction policy, optimal allocation of labor force and intelligence of manufacturing industry. *China Industrial Economy*, (05), 116-134.
- [10] Westover, J., & Tang, F. (2025). The invisible AI workforce: Redefining technical talent acquisition for artificial intelligence. Preprints 2025, 2025091488.

- [11] Ma, B. C., & Liao, K. Z. (2025). How does artificial intelligence application affect firm innovation: Evidence from online recruitment. *Journal of Zhongnan University of Economics and Law*, (06), 132-145. <https://doi.org/10.19639/j.cnki.issn1003-5230.2025.0061>.
- [12] Lu, D., Yu, D., Huang, D., & Yang, R. M. (2020). Internal training or external recruitment: Which one better promotes firm innovation? *China Industrial Economy*, (10), 157-174. <https://doi.org/10.19581/j.cnki.ciejournal.2020.10.007>.
- [13] Wang, Z. Q., Wang, Y. Z., & Wang, S. H. (2019). An analysis of the real level of capital efficiency and financial risk in China's real economy: What is the crux of the low efficiency and level of financial services for the real economy? *Management World*, 35(2), 19.
- [14] Ju, X. S., Lu, D., & Yu, Y. H. (2013). Financing constraints, working capital management and the sustainability of firm innovation. *Economic Research Journal*, 48(01), 4-16.
- [15] Yavuz, M. S., & Çalık, H. (2025). AI and ML patent intensity and firm performance: A machine learning-based lagged analysis. ResearchGate.
- [16] Huang, Q. H., Yu, Y. Z., & Zhang, S. L. (2019). Internet development and manufacturing productivity improvement: Internal mechanism and Chinese experience. *China Industrial Economy*, (08), 5-23. <https://doi.org/10.19581/j.cnki.ciejournal.2019.08.001>.
- [17] Chen, X., Wu, Y. J., & Zhu, J. H. (2024). Local talent introduction and total factor productivity from the perspective of new quality productive forces. *Journal of Business Economics and Management*, 46(12), 104-120. <https://doi.org/10.19616/j.cnki.bmj.2024.12.006>.
- [18] Li, L., & Wang, T. Y. (2023). "Peacocks flying southeast": High-quality economic development and talent flow. *The Journal of Quantitative & Technological Economics*, 40(02), 5-24. <https://doi.org/10.13653/j.cnki.jqte.2023.02.001>.
- [19] Mao, F. F., & Zheng, F. (2021). How do talent introduction policies affect the labor market?. *Journal of Business Economics and Management*, (11), 62-77. <https://doi.org/10.14134/j.cnki.cn33-1336/f.2021.11.005>.
- [20] Sun, R. & Sun, Y. J. (2021). Quantitative research on local innovation and entrepreneurship talent introduction policies in China. *Science Research Management*, 42(06), 29-44. <https://doi.org/10.20201/j.cnki.ssstm.2021.06.003>.
- [21] Zhang, T. B., & Wang, L. (2024). Policy signals and long-term residence intention of floating population. *The Journal of World Economy*, 47(07), 97-122. <https://doi.org/10.19985/j.cnki.cassjwe.2024.07.003>.
- [22] Sun, W. Y., & Liu, Y. S. (2023). Research on the impact mechanism of artificial intelligence on the labor market. *East China Economic Management*, 37(03), 1-9. <https://doi.org/10.19629/j.cnki.34-1014/f.220706008>.
- [23] Hadlock, C. J., & Pierce, J. R. (2010). New evidence on measuring financial constraints: Moving beyond the KZ index. *Review of Financial Studies*, 23(5), 1909-1940.
- [24] Du, Y. H., Zhao, R., & Zhao, X. Y. (2024). The contribution of higher education to innovation economy: An empirical study based on firm panel data. *Peking University Education Review*, 22(01), 18-40+187-188.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

A Study on the Impact of ESG Performance on Corporate Performance—Based on the Moderating Effect of Digital Transformation

Qian Li*

School of Statistics and Mathematics, Yunnan University of Finance and Economics, Kunming, China

**Corresponding author: Qian Li.*

Abstract

Against the backdrop of the parallel development of the “dual-carbon” strategy and the digital economy, the impact of ESG performance on corporate performance and the moderating effect of digital transformation have become valuable research topics. This paper uses A-share listed companies from 2016 to 2024 as a sample to explore the impact of ESG performance on corporate performance (ROA) and the moderating effect of digital transformation. The study finds that ESG performance has a significant negative impact on current performance, exhibiting obvious “short-term pain” characteristics and potential long-term positive effects that have not yet been fully realized. The combined costs of digital transformation and ESG construction exert a negative moderating effect. Heterogeneity analysis shows that the pollution attributes of an industry weaken the promoting effect of ESG performance on corporate performance, with heavily polluting industries showing even lower marginal effects. Furthermore, factors such as company size and growth significantly positively impact corporate performance, while financial leverage and the short-term costs of digitalization have a negative impact. Based on this, this paper proposes countermeasures from the perspectives of government, enterprises, and investment, providing theoretical reference and practical guidance for ESG-enabled corporate performance transformation and upgrading.

Keywords

ESG performance, corporate performance, digital transformation, moderating effect, industry heterogeneity

1. Introduction

Green development is a new form of economic growth and social development with efficiency, harmony and sustainability as its goals [1]. In 2020, China clearly proposed the “dual carbon” goal of “striving to achieve carbon peak before 2030 and carbon neutrality before 2060”, marking the entry of the country's ecological civilization construction into a new stage. Against this background, enterprises, as carriers of technological innovation and low-carbon practices, are particularly important for their green and sustainable development. The ESG (Environment, Society and Governance) concept has received much attention because it is highly compatible with the “dual carbon” goal. At the same time, the digital economy, as a new economic form with data resources as the key element, modern information networks as the main carrier, and the integration and application of information and communication technologies as the driving force, is providing enterprises with

a large number of innovation opportunities through big data, artificial intelligence and other technologies, and promoting enterprises to maintain competitiveness in the rapidly changing market environment [2]. Digital transformation not only provides enterprises with technical support for innovative development, but also plays an irreplaceable role in resource allocation, efficiency improvement and risk management. Against the backdrop of the coordinated advancement of the “dual-carbon” strategy and the digital economy, how ESG performance affects corporate performance, and whether digital transformation will alter this mechanism, have become important topics requiring in-depth exploration.

Corporate performance is a core indicator used to evaluate the effectiveness of corporate strategy execution and the efficiency of resource utilization. It reflects a company's capabilities in market competition and supports decision-making and improvement to achieve long-term sustainable development. Since the reform and opening up, the performance of Chinese enterprises has been steadily improving, but many problems remain, such as weak performance management awareness, outdated performance appraisal data management methods, and a lack of scientific rigor in content and indicator design. These problems continue to affect corporate resilience and its domestic and international competitiveness. At the academic research level, although existing literature has confirmed that good ESG performance can improve corporate performance, there are still disagreements and blind spots regarding the internal mechanisms by which ESG performance is transformed into financial performance. Some studies have found that ESG investments are difficult to translate into financial performance in the short term, and may even have an inhibitory effect on performance due to increased costs; other studies have pointed out that there are significant differences in the efficiency of ESG performance transformation among different companies, and the root causes of these differences have not yet been fully explained. Crucially, most existing research focuses on the independent impacts of ESG performance or digital transformation on corporate performance, neglecting their interaction. Whether digital transformation can act as a catalyst or obstacle to the conversion of ESG performance into corporate performance remains unclear, and its moderating mechanism has not been systematically revealed. This research gap provides a space for exploration in this paper.

Currently, the criticality of ESG performance in the field of corporate performance is widely recognized, but the following shortcomings still exist, urgently requiring further research: Firstly, the breadth of research is insufficient; most literature focuses on the direct impact of ESG on corporate performance, lacking attention to influencing factors and boundary conditions. Furthermore, there is a lack of research combining ESG performance with digital transformation to explore pathways through which “digital value transformation” acts as a moderating mechanism for corporate performance. In the current context of vigorously advocating green development and the digital economy, digital transformation, as a crucial technological variable in corporate operations, is highly likely to modulate the relationship between ESG performance and corporate performance by altering information transmission rates and enhancing resource integration capabilities. Therefore, in-depth exploration of the synergistic or competitive effects of ESG performance and digital transformation in the field of corporate performance has significant practical implications. On the other hand, research depth is lacking. Existing studies on the relationship between ESG and corporate performance fail to fully explain the contradiction that “some companies' ESG investments are efficiently converted into performance, while others see minimal results.” Digital transformation, as a key means for companies to improve operational efficiency and optimize governance structure, largely addresses this issue, but the inherent logic, direction, and intensity of its moderating effect still require empirical verification. Therefore, this paper aims to use digital transformation as a moderating variable, analyzing relevant statistical indicators to explore the moderating mechanism of corporate digital transformation in the process of ESG performance affecting corporate performance from both theoretical and empirical perspectives. This provides theoretical support and practical guidance for companies to achieve performance upgrades through the interaction between ESG and digital transformation.

This research is of significant importance. From a theoretical perspective, it enriches the research perspective on the relationship between ESG performance and corporate performance. This paper breaks through the existing research's singular focus on the direct relationship between the two, introducing digital transformation as a moderating variable to reveal the conditional factors that influence corporate performance through ESG, thus improving the theoretical system of ESG performance effects; it also deepens the research on the economic effects of digital transformation. Existing research largely focuses on the direct impact of digital transformation on corporate performance. This paper places it within the framework of the relationship

between ESG and corporate performance, exploring its moderating role and expanding the application scenarios and theoretical boundaries of digital transformation research; it also discovers the inherent logic of ESG performance transformation. By analyzing the moderating effect of digital transformation, it explains the root causes of differences in the efficiency of ESG investment performance transformation among different companies, providing a new analytical dimension for solving difficulties in academic research. From a practical perspective, it provides development direction for enterprises, guiding them to dynamically adjust their ESG investment strategies according to their own level of digital transformation during the ESG construction process. Simultaneously, it reveals the technology spillover effects driven by internal governance changes or compliance triggered by “mandatory disclosure,” providing new ideas for sustainable corporate development; it provides reference for policymakers, offering empirical evidence for the government to formulate ESG-related policies and digital transformation support policies, helping to build a policy environment for the real-time linkage of “ESG + digitalization,” and promoting high-quality corporate development; and it provides decision-making basis for investors, helping them to more comprehensively evaluate the value of corporate ESG investment, pay attention to the impact of digital transformation level on ESG performance effects, and improve the scientific nature of investment decisions.

In terms of research methodology, this paper adopts a combination of normative and empirical research methods. The normative research is reflected in the systematic review of domestic and international literature related to ESG performance, digital transformation, and corporate performance. Based on theories such as information asymmetry and resource-based theory, it constructs a theoretical framework for how digital transformation moderates the impact of ESG performance on corporate performance and proposes research hypotheses. The empirical research selects A-share listed companies from 2016 to 2024 as research samples. Through the construction of a multiple linear regression model, it empirically tests the direct impact of ESG performance on corporate performance and the moderating effect of digital transformation. Robustness tests and heterogeneity analysis ensure the reliability of the research conclusions.

2. Literature Review

2.1 The Impact of ESG Performance on Corporate Performance

The impact of ESG performance on corporate performance is an important topic in academia.

Some studies support that ESG has a positive driving effect on corporate performance. Hu and Qin found, based on data from Shanghai and Shenzhen A-shares from 2013 to 2023, that ESG performance can improve corporate performance by alleviating financing constraints and reducing agency costs [3]; Shi used A-share listed companies from 2018 to 2022 as a sample to confirm that innovation capability plays a mediating role in the positive relationship between ESG performance and corporate performance [4]; Zhou et al. also verified that ESG rating has a positive impact on corporate performance [5]; Tao et al. used data from A-share agricultural enterprises from 2013 to 2023 to study and found that green technology innovation plays a significant moderating role in the relationship between ESG performance and corporate performance, which helps to enhance the positive impact of ESG performance on corporate performance [6].

Another view is that ESG investment may have an inhibitory effect on performance in the short term due to increased corporate costs. The “trade-off hypothesis” proposed by Preston and O'Bannon laid the theoretical foundation for this view [7]. They believed that corporate resources are limited and that expenditures on social responsibility and environmental governance will directly squeeze profit margins, leading to a decline in short-term financial performance. Subsequently, Margolis et al. further confirmed through a large-scale meta-analysis that the positive correlation between corporate social responsibility and financial performance usually takes a long time to manifest [8]. In the early stage of investment, due to the cost first and the benefit lagging, negative short-term effects are often observed. In China, Zhang and Zhao found through analysis of Chinese A-share listed companies that ESG performance is negatively correlated with corporate financial performance (ROA, ROE) in the short term [9]. Gao et al. revealed that when there is a large divergence in the market's ESG rating of enterprises, it will increase the uncertainty of corporate information and financing costs, thereby having a negative impact on the total factor productivity of enterprises [10]. Xie and Lü also confirmed this rule, that enterprises will increase compliance costs in the short term when fulfilling ESG responsibilities, but in the long term it will help improve the performance of foreign investment [11]. These studies collectively

reveal the “short-term pain” characteristic of ESG investment, namely, the early investment of enterprises in environmental, social and governance aspects will significantly increase operating costs, while the corresponding long-term benefits such as cost reduction and efficiency improvement, risk management and competitiveness reconstruction will take a long time to realize.

2.2 The Moderating Role of Digital Transformation

The relationship between enterprise digital transformation and performance and value presents a “double-edged sword” situation. On the one hand, digital transformation can promote the improvement of enterprise performance and the reconstruction of value. Wu et al. proposed that digital technology can optimize production processes, reduce communication costs and improve resource allocation efficiency [12]; Bharadwaj showed that as the digital capabilities of enterprises mature, innovative models can be built to form competitive advantages, thereby achieving long-term value growth [13]. On the other hand, digital transformation is very likely to have a negative effect and inhibitory effect on performance in the short term due to the high cost investment. Brynjolfsson and McAfee pointed out that initial costs will squeeze profit margins and drag down productivity growth in the short term, forming a “digital paradox” [14]; Liu et al. empirically found that digital transformation and corporate performance have a U-shaped relationship [15]. The underlying mechanism is that high initial investment leads to reduced performance, and performance increases significantly after crossing the critical point. In summary, digital transformation has the characteristics of “accumulation and release”. Xiao et al. found that digital transformation indirectly affects corporate performance by optimizing human capital structure, but its positive effect requires a long adjustment period to manifest [16].

This paper focuses on the moderating role of digital transformation in the relationship between ESG performance and corporate performance. Its internal logic chain lies in the bidirectional transformation from compliance-driven technology spillover to digital value [17]: First, in order to achieve ESG goals such as carbon emission monitoring and resource management, enterprises need to build digital infrastructure. Although this initial investment stems from the pressure of compliance disclosure, its effectiveness will spread to the entire process of enterprise operation and management, i.e. the entire business chain, through optimizing production processes, thereby realizing the transition from “external regulatory pressure” to “internal business motivation”. Secondly, digital transformation utilizes big data and cloud computing technologies, which can affect the performance conversion efficiency of ESG investment in different situations. Highly digital enterprises can accelerate the realization of ESG value through information transmission, resource integration and risk control mechanisms, while lowly digital enterprises may fall into a short-term performance dilemma due to dual cost pressures [18].

3. Theoretical Analysis and Research Hypotheses

3.1 Theoretical Analysis of the Impact of ESG Performance on Corporate Performance

Based on the resource-based theory and the trade-off hypothesis [19-20], corporate investment in environmental, social, and governance (ESG) typically exhibits a “cost-first, benefit-later” characteristic. First, according to the trade-off hypothesis, corporate resources are limited in the short term. To improve ESG performance, companies need to invest heavily in green technology research and development, pollution control facilities, and employee welfare improvements. These expenditures are often included in current costs in financial accounting, directly squeezing corporate profit margins and causing a decline in performance indicators such as return on total assets (ROA) in the short term, resulting in obvious “short-term pain.” Second, according to the information asymmetry theory [21-22], the non-financial benefits brought by ESG investment (such as enhanced brand reputation and reduced financing constraints) have strong positive externalities and uncertainties [23]. These benefits often need a long time window for the market and investors to gradually digest and reflect them in corporate performance. Therefore, in the initial stages of investment, high compliance and operational costs dominate, while long-term benefits have not yet materialized. Based on this, this paper proposes the following hypothesis:

H1: ESG performance has a lagged positive impact on corporate performance.

3.2 The Moderating Logic of Digital Transformation in the Relationship between ESG Performance and Corporate Performance

As a key technological variable in corporate operations, digital transformation and ESG construction exhibit a “double cost” logic in their interaction. On the one hand, digital transformation itself is a high-investment, high-risk project. According to the “digital paradox” theory, companies need to build expensive digital infrastructure and introduce technologies such as big data and artificial intelligence in the early stages of transformation, which will generate huge financial expenditures in the short term, squeezing current profits. On the other hand, when companies simultaneously promote ESG construction and digital transformation, resource competition and cost superposition effects will occur. Companies with high levels of digitalization often need to invest in more cutting-edge digital technologies to achieve ESG goals such as accurate carbon emission monitoring and transparent supply chain management, which further increases their financial burden. Companies with low levels of digitalization may lack effective resource integration capabilities and fall into a passive investment dilemma when dealing with ESG regulatory pressure, making the double cost pressure even more prominent. Therefore, the cost effects of digitalization may overshadow its enabling effects in the short term, but it possesses potential long-term benefits. Based on the above analysis, this paper proposes the following hypothesis:

H2: Digital transformation plays a significant moderating role in the relationship between ESG performance and corporate performance; that is, the level of digitalization alters the direction and intensity of the impact of ESG performance on corporate performance.

4. Research Design

4.1 Sample and Data Sources

This paper selects A-share listed companies data from 2016 to 2024 as the research sample, and performs the following screening on the initial data: (1) Excludes abnormal operating samples such as ST, *ST, and PT; (2) Excludes samples from the financial and insurance industries, as the characteristics and regulatory requirements of this type of industry differ greatly from other industries, making horizontal comparison difficult; (3) Excludes samples with missing variables to ensure data integrity; (4) Performs a 1% winsorize (tailing) on all continuous variables to eliminate the influence of extreme values. After screening, 26,121 sample observations were finally obtained. The corporate ESG performance data comes from Huazheng Index, a third-party authoritative rating agency published by Wind. The data related to digital transformation are constructed using Python text analysis technology, and other control data are all from the CSMAR database.

4.2 Variable Definitions

Explicit Variable: Corporate Performance (ROA). In existing studies, scholars mostly quantify corporate performance from a financial perspective. The indicators, including Return on Assets (ROA), Return on Equity (ROE), Tobin's Q, and Earnings Per Share (EPS), provide quantitative accordance from the perspectives of asset efficiency, capital return, market expectations, and shareholder interests. This paper aims to explore whether ESG performance can be transformed into a company's core competitive advantage and generate additional revenue, thereby improving performance. Based on Shi Yiling's methodology, this paper uses ROA as the performance evaluation standard. It not only directly reflects the profitability of assets but also embodies the company's comprehensive competence in resource utilization and value creation, demonstrating the company's operational development and resource value to a certain extent, making it highly representative.

Explanatory Variable: ESG Performance. This paper uses annual data from Huazheng ESG Ratings. Huazheng ESG Ratings, from highest to lowest, are AAA, AA, A, BBB, BB, B, CCC, CC, and C. This paper assigns scores from 1 to 9 based on these ratings, with higher scores indicating better ESG performance.

Mediating Variable: Digital Transformation. This paper draws on the research method of Wu et al. to construct enterprise digital transformation indicators through text recognition of listed companies' annual reports [24]. Compared with a single financial input indicator, text analysis can capture the management's emphasis on digitalization and actual application scenarios from the strategic perspective of the enterprise, and is more comprehensive and forward-looking. Moreover, the indicators cover keywords of underlying

technologies and practical applications, which can depict the degree of penetration of enterprises in digital transformation and effectively measure the substantial progress of enterprise digitalization. More convincingly, the thesaurus constructed by Wu et al. has been widely used in Chinese core journals, and its reliability and validity have been fully verified, which helps to improve the objectivity and comparability of the empirical analysis results in this paper. The specific method is as follows: using Python text analysis technology, keyword retrieval is performed on the “Management Discussion and Analysis” (MD&A) section of the annual reports of A-share listed companies. The retrieval scope covers five dimensions: artificial intelligence, big data, cloud computing, blockchain and digital technology applications. The frequency of occurrence of keywords in the above five dimensions is summed to obtain the original total word frequency of enterprise digital transformation. Considering the obvious right-skewed characteristic of word frequency data, this paper performs natural logarithmic transformation on the total word frequency after adding 1, i.e., $Digital = \ln(1 + Total\_Keywords)$. A higher score indicates a deeper level of digital transformation for the enterprise.

Interaction Term: ESG Performance×Digital Transformation (ESG×Digital). To test the moderating effect of digital transformation, this paper constructs the interaction term ESG×Digital, which is the core variable for testing hypothesis H2. If the coefficient of the interaction term is significantly positive, it indicates that digital transformation strengthens the positive impact of ESG performance on corporate performance; conversely, if the coefficient is significantly negative, it indicates that digital transformation weakens this positive impact.

Control Variables: This paper selects the following control variables: firm size (Asset), financial leverage (Debt), cash flow ratio (Cashflow), revenue growth rate (Growth), number of directors (Boardsize), duality (Duality), and government subsidies (InSubsidy). Additionally, a year dummy variable (Year) and a firm fixed-effects dummy variable (Firm) are included. The definitions of these variables are shown in Tab.1.

Table 1: Variable definition

Variable Type	Variable Name	Variable Symbol	Variable Description
Explanatory variable	Return on Total Assets	ROA	Year-end Net Profit / Total Assets
Explanatory variables	ESG performance	ESG	Huazheng ESG Rating Results
Mediator variables	Digital transformation level	Digital	The natural logarithm of the total frequency of words in the five dimensions of artificial intelligence, big data, cloud computing, blockchain, and digital technology applications plus 1.
control variables	Company Size	Asset	Natural logarithm of total annual assets
	Financial Leverage	Debt	Total liabilities / Total assets
	Cash Flow Ratio	Cashflow	Net cash flow from operating activities / Total assets
	Revenue Growth	Growth	Year-end revenue growth / Beginning-of-year total revenue
	Number of Directors	Boardsize	The natural logarithm of the number of board members
	Dual Roles (Both Director and Director)	Duality	If the chairman and the general manager are the same person, assign a value of 1; otherwise, assign a value of 0.
	Government Subsidies	InSubsidy	External policy impact
	Company Year	Firm	Company fixed-effects dummy variable
	Company Size	Year	Annual dummy variable

Source: Author's survey

4.3 Model Construction

To study the impact of ESG performance on corporate performance, and to verify Hypothesis 1, a regression model (1) was constructed:

$$ROA_{i,t} = \alpha_0 + \beta_1 ESG_{i,t} + \sum Control_{i,t} + \sum Firm + \sum Year + \varepsilon_{i,t} \quad (1)$$

Where, the subscript i represents the individual, t represents the year, $\sum Control_{i,t}$ is the set of control variables, Year and Firm are the year fixed effect and company fixed effect, respectively (industry attributes

are inherent characteristics of enterprises that do not change over time, and the company fixed effect will completely absorb the industry fixed effect), and ε is the residual value.

Furthermore, to explore the moderating effect of digital transformation, and to verify Hypothesis 2, a moderating variable Digital and an interaction term ESG×Digital were added to model (1) to construct model (2):

$$ROA_{i,t} = \alpha_0 + \beta_2 ESG_{i,t} + \beta_3 Digital_{i,t} + \beta_4 (ESG_{i,t} \times Digital_{i,t}) + \sum Control_{i,t} + \sum Firm + \sum Year + \varepsilon_{i,t} (2)$$

5. Empirical Results and Analysis

5.1 Descriptive Statistics

To initially observe the overall characteristics of the sample, this paper conducted descriptive statistical analysis on the main variables. Statistical indicators included sample size, mean, standard deviation, median, minimum, and maximum values. The results are shown in Table 2. Fig.1 to Fig.3 visually illustrate the distribution density of the core variables ROA, ESG, and Digital, facilitating understanding of the sample characteristics. Overall, there are certain differences among the variables, which can well reflect the heterogeneity of the observed enterprises in terms of operating conditions and governance levels.

From the perspective of the explained variable, the mean of firm performance (ROA) is 0.0321, the standard deviation is 0.0680, the minimum is -0.2479, the maximum is 0.2145, and the median is 0.0337. It can be seen that there are significant differences in profitability among different enterprises, with some enterprises experiencing negative returns, indicating that the distribution of firm performance has a certain degree of dispersion.

From the core explanatory variables, the mean of ESG performance (ESG) is 4.116, the median is 4, the minimum is 1, the maximum is 7, and the standard deviation is 1.165. Overall, ESG performance is at a moderate level, but there are significant differences between companies, indicating uneven practices in environmental, social responsibility, and corporate governance.

From the control variables, the mean of firm size (Asset) is 22.413, the median is 22.24, the minimum is 20.2, the maximum is 26.21, and the standard deviation is 1.230, showing that while the overall size of firms is not significantly different, some differentiation exists. The mean of the debt-to-equity ratio (Debt) is 0.426, the median is 0.417, the minimum is 0.068, the maximum is 0.928, and the standard deviation is 0.200, indicating that the overall leverage level is within a moderate range, but there are significant differences in financial structures among different companies. The mean of the cash flow ratio is 0.0503, the minimum is -0.1365, the maximum is 0.2411, and the standard deviation is 0.0649, indicating that some companies have negative operating cash flow and that there are significant differences in operating capabilities among companies. The mean of growth is 0.2874, but the standard deviation reaches 0.7372, with a minimum of -0.6840 and a maximum of 4.8252, indicating high disparity in growth capabilities among companies. Some companies are in a phase of rapid growth, while others are experiencing stagnant or even declining growth. The mean of dual roles (chairman and general manager) is 0.3088, indicating that approximately 30% of companies have the same person holding both the chairman and general manager positions. The mean of the government subsidy variable (lnSubsidy) is 15.9478, the minimum is 8.5174, the maximum is 20.1627, and the standard deviation is 2.0480, reflecting the varying degrees of government subsidies received by companies.

Furthermore, the mediating variable “Digital Transformation Level” has a mean of 1.331, a minimum of 0, a maximum of 4.727, and a standard deviation of 1.305, revealing a uneven progress in digitalization. Some companies have not yet significantly implemented digital transformation, while others have a high level of digitalization.

Overall, the descriptive statistical analysis results obtained in this paper generally reflect the actual situation of listed companies and possess a certain degree of reliability.

Table 2: Descriptive statistics

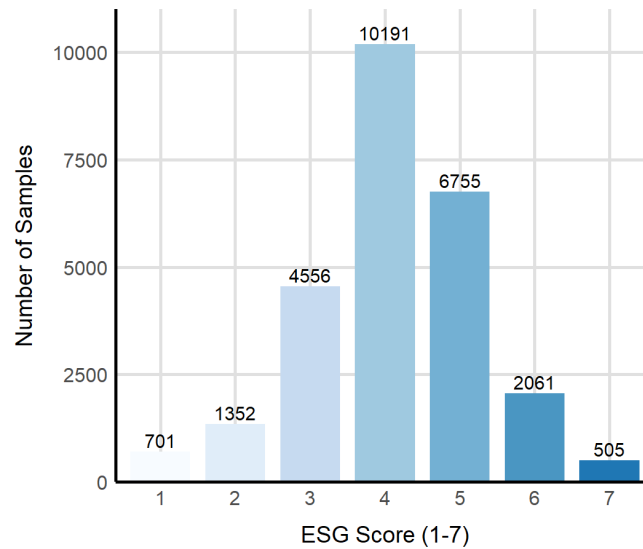
Variable	Obs	Mean	Sd	Min	Md	Max
ROA	26121	0.032	0.068	-0.248	0.034	0.214
ESG	26121	4.116	1.165	1	4	7

Variable	Obs	Mean	Sd	Min	Md	Max
Asset	26121	22.413	1.230	20.200	22.240	26.210
Debt	26121	0.426	0.200	0.068	0.417	0.928
Cashflow	26121	0.050	0.065	-0.136	0.048	0.241
Growth	26121	0.287	0.737	-0.684	0.110	4.825
Duality	26121	0.309	0.462	0	0	1
lnSubsidy	26121	15.948	2.048	8.517	16.214	20.163
Digital	26121	1.331	1.305	0	1.099	4.727

Note: Obs is the sample size, Mean is the sample mean, Sd is the sample standard deviation, Min is the sample minimum, Md is the sample median, and Max is the sample maximum.

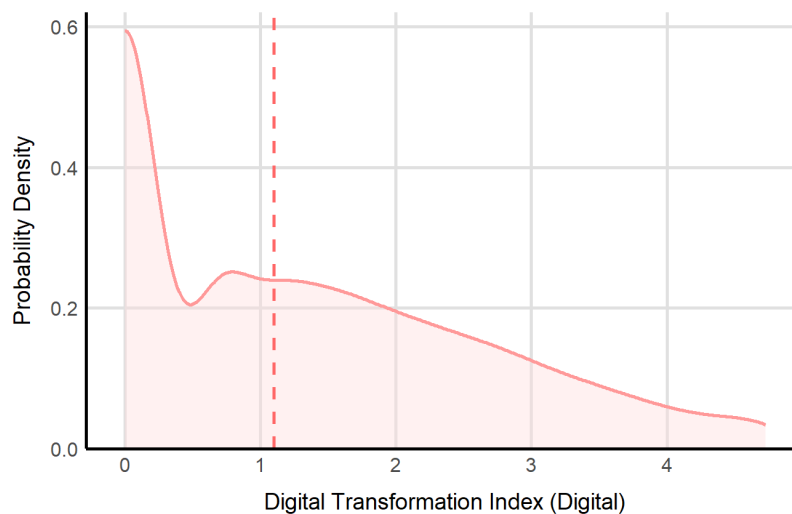
Source: Author's survey

Figure 1: ESG distribution bar chart
Distribution of ESG Rating Scores



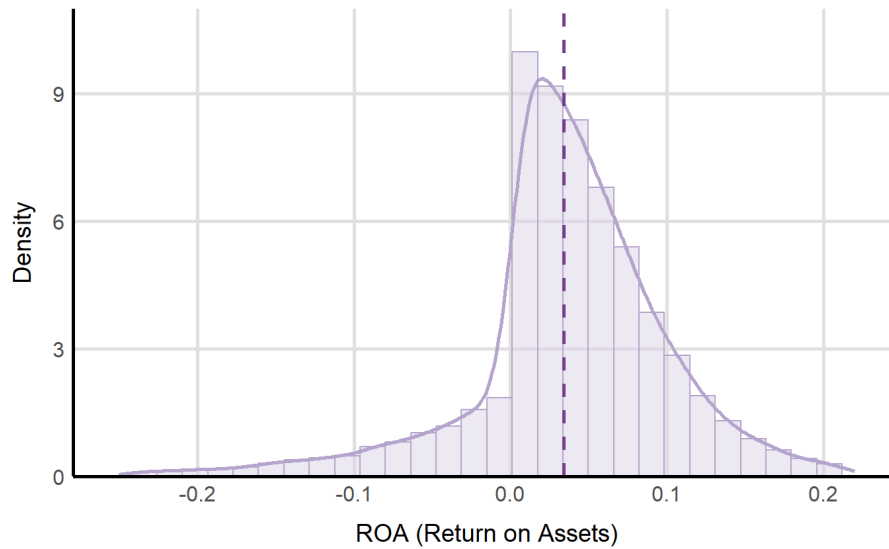
Source: Author's survey.

Figure 2: Digital Distribution Density Map
Distribution Characteristics of Enterprise Digital Transformation Level
Dashed line represents industry median



Source: Author's survey

Figure 3: ROA distribution histogram
Distribution of Corporate Performance (ROA) Samples
 Dashed line represents the mean



Source: Author's survey.

5.2 Correlation Analysis

The results of the correlation analysis are shown in Table 3. The correlation coefficients among different variables were generally below 0.500. The correlation coefficient between ESG performance (ESG) and corporate performance (ROA) was 0.221, significantly positive at the 1% level, indicating that companies with better ESG performance typically have higher operational performance. Meanwhile, there was a negative correlation between corporate digitalization level and corporate performance (ROA), with a correlation coefficient of -0.0612. Furthermore, there was also a certain degree of correlation between ESG performance (ESG) and corporate digitalization level (Digital). It is worth noting that corporate digital transformation often requires high upfront investment, which may have a certain impact on corporate profitability in the short term. Therefore, it showed a negative correlation with corporate performance in the correlation analysis, but this result still needs further verification through regression analysis after controlling for relevant variables.

To test for the existence of multicollinearity in the model, this paper conducted a variance inflation factor (VIF) test on each explanatory variable. The results are shown in Table 4. The VIF values of all variables are significantly less than 10, with the maximum value being 1.62 and the average VIF being 1.20. This indicates that there is no significant multicollinearity problem among the explanatory variables in the model, and the model specification has good stability.

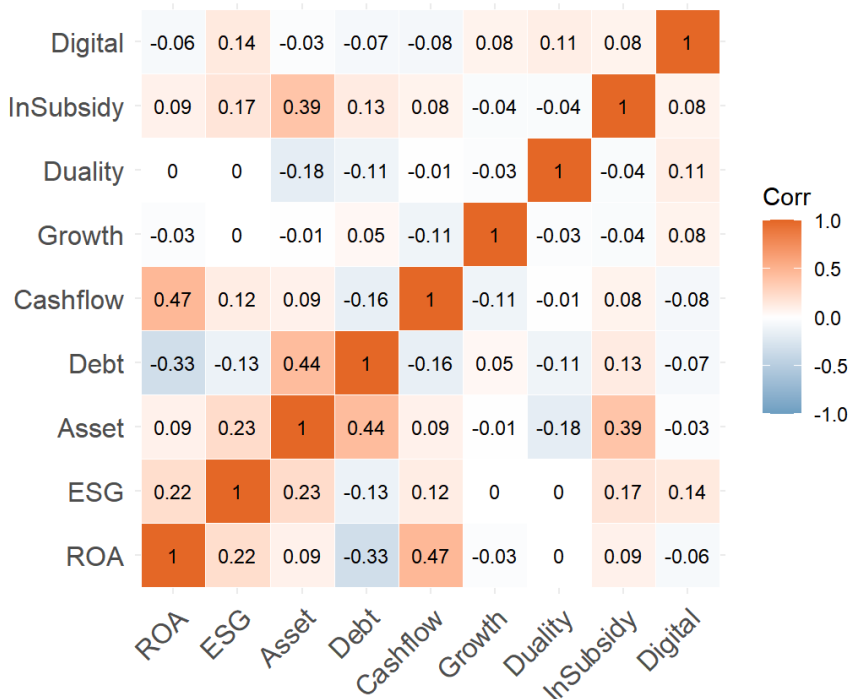
Table 3: Correlation analysis results

Variable	ROA	ESG	Asset	Debt	Cashflow	Growth	Duality	lnSubsidy	Digital
ROA	1								
ESG	0.2206***	1							
Asset	0.0850***	0.2276***	1						
Debt	-0.3324***	-0.1308***	0.4399***	1					
Cashflow	0.4663***	0.1196***	0.0888***	-0.1648***	1				
Growth	-0.0265*	0.0022	-0.0133*	0.0484***	-0.1134***	1			
Duality	0.0026	0.0023	-0.1757***	-0.1067***	-0.0091	-0.0258***	1		
lnSubsidy	0.0870***	0.1674***	0.3851***	0.1277***	0.0804***	-0.0428***	-0.0382***	1	
Digital	-0.0612***	0.1372***	-0.0328***	-0.0720***	-0.0838***	0.0759***	0.1115***	0.0840***	1

Note: *, **, and *** indicate significance at the 10%, 5%, and 1% significance levels, respectively.

source: Author's survey

Figure 4: Correlation heatmap
Correlation Matrix of Research Variables



Source: Author's survey

Table 4: VIF test results

variable	VIF	1/VIF
Asset	1.62	0.616
Debt	1.40	0.714
InSubsidy	1.20	0.834
ESG	1.16	0.858
Cashflow	1.09	0.918
Digital	1.06	0.942
Duality	1.05	0.955
Growth	1.02	0.978
Mean VIF	1.20	

Source: Author's survey

5.3 Regression Analysis

5.3.1 Impact of ESG Performance on Corporate Performance

To examine the direct impact of ESG performance on corporate performance, this paper constructs model (1) for regression analysis. Based on controlling for firm size (Asset), financial leverage (Debt), cash flow ratio (Cashflow), revenue growth rate (Growth), duality (Duality), government subsidies (InSubsidy), and year fixed effects, a firm fixed effects model is adopted, and the standard error is adjusted for clustering at the firm level to mitigate the impact of heteroscedasticity and serial correlation on the estimation results. The regression results are shown in Table 5.

Column (1) of Table 5 reports the baseline regression results of ESG performance on corporate performance (ROA). The regression coefficient of the core explanatory variable ESG is -0.0009, and it is significant at the 5% level ($t = -2.16$, $p = 0.031$). This result indicates that during the sample period, there is a significant negative relationship between corporate ESG performance and current corporate performance; that is, the better the ESG performance of a company, the lower its current financial performance. This result aligns with the “short-term pain” phenomenon described in the preceding literature: firms often need to pay upfront costs for environmental, social, and governance (ESG) investments, while the resulting improvements in operational

efficiency and risk reduction have a certain lag, potentially crowding out financial performance in the short term [7,9]. This also means that the performance transformation of ESG is not immediate but requires a certain time frame.

From the perspective of control variables, the coefficient for firm size (Asset) is significantly positive (0.0280, $p < 0.01$), indicating that larger firms have stronger profitability and resource integration capabilities. The coefficient for the debt-to-equity ratio (Debt) is significantly negative (-0.1862, $p < 0.01$), suggesting that excessive financial leverage weakens firm performance. The cash flow ratio (Cashflow) is significantly positive (0.2371, $p < 0.01$), confirming that robust cash flow is a crucial guarantee for firm performance. The revenue growth rate (Growth) is significantly positive (0.0062, $p < 0.01$), showing that firms with higher growth rates perform better. The coefficients for duality and government subsidies (lnSubsidy) are negative but not significant, indicating that their direct impact on current performance is limited. A significant year-fixed effect indicates that the macroeconomic environment in different years has a systematic impact on firm performance.

The above results provide the necessary preconditions for subsequent testing of the lagged effects of ESG performance: if ESG investment does indeed have the characteristics of cost-first and benefit-later, then the lagged term of introducing ESG should show a significant positive impact. Therefore, this paper will further discuss the dynamic effects of ESG performance to fully reveal its true impact path on corporate performance.

Table 5: ESG performance benchmark regression results on firm performance

Variable	Coefficient	Standard error	t-value	P> t	95% confidence interval
ESG	-0.0009**	0.0004	-2.16	0.031	[-0.0018, -0.0001]
Asset	0.0280***	0.0017	16.66	0.000	[0.0247, 0.0313]
Debt	-0.1862***	0.0065	-28.70	0.000	[-0.1989, -0.1735]
Cashflow	0.2371***	0.0102	23.16	0.000	[0.2170, 0.2571]
Growth	0.0062***	0.0007	8.57	0.000	[0.0048, 0.0076]
Duality	-0.0006	0.0015	-0.36	0.718	[-0.0036, 0.0025]
lnSubsidy	-0.0002	0.0002	-1.10	0.272	[-0.0006, 0.0002]
Constant term	-0.5047***	0.0358	-14.08	0.000	[-0.5750, -0.4345]
Year fixed effects	control				
Company fixed effects	control				
Sample size	26,121				
Number of firms	4,531				
R ² (within)	0.2061				

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively; standard error is adjusted for clustering at the company level.

Source: Author's survey

5.3.2 Moderating Effect Test of Digital Transformation

To test the moderating effect of digital transformation on the impact of ESG performance on corporate performance, this paper introduces digital transformation (Digital) and its interaction term with ESG performance (ESG×Digital) on the basis of model (1), constructing model (2). To avoid multicollinearity, ESG and Digital were centered to generate the interaction term. The regression results are shown in Table 6 and Figure 6.

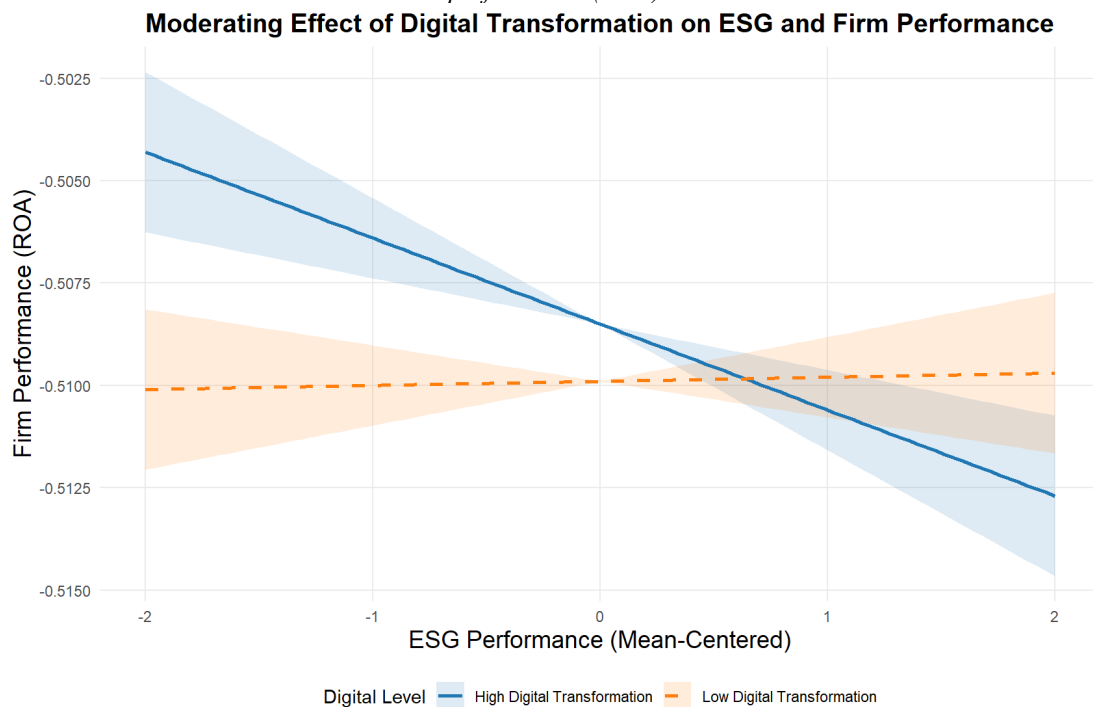
Column (2) of Table 6 reports the estimation results of the moderating effect. The coefficient of the core explanatory variable ESG_c is -0.0010, which is still significant at the 5% level ($t = -2.24$, $p = 0.025$), consistent with the baseline results. The coefficient of digital transformation (Digital_c) is 0.0007, but it did not pass the significance test ($p = 0.284$), indicating that the direct impact of the current degree of digital transformation on corporate performance is not significant, which may be due to the long input-output cycle of digital construction. More importantly, the coefficient of the interaction term ESG_x_Digital is -0.0011 and is significant at the 1% level ($t = -3.17$, $p = 0.002$). This result indicates that digital transformation significantly and negatively moderates the relationship between ESG performance and corporate performance. In other words, for companies with a higher degree of digital transformation, the negative effect of ESG performance on current performance is more pronounced.

This result is consistent with hypothesis H2 (digital transformation plays a moderating role), but the moderating direction differs from the expected positive one. A possible reason is that digital transformation itself requires significant upfront investment. When companies simultaneously promote ESG development and digital innovation, the dual cost pressures may overlap in the short term, further squeezing profit margins and thus reinforcing the short-term negative effects of ESG. Furthermore, highly digitalized companies often face more complex technological challenges, and the integration of ESG concepts may require a longer adjustment period, leading to a temporary decrease in performance conversion efficiency. This finding reveals that digital transformation can become a “stumbling block” to ESG performance conversion at specific stages and for different companies, providing a new perspective for understanding the combined effects of corporate ESG and digitalization.

From the perspective of control variables, the signs and significance of the coefficients of each variable remained relatively stable, indicating that the model setting was relatively robust. The year fixed effect remained significant, indicating that the impact of the macro environment on performance continued.

Further understanding, referring to the research of Zhao et al., digital transformation has a significant threshold effect on the improvement of total factor productivity [25]. that is, only when the investment in digitalization exceeds a certain threshold can the production efficiency dividend be fully released. The negative adjustment effect found in this paper is essentially caused by the superposition of the dual costs of ESG construction and digital transformation in the early stage of investment. When enterprises simultaneously promote two high-investment, long-cycle strategic changes, the resource competition effect doubles the short-term cost pressure. This finding also implies the possible dynamic evolution law of ESG and digitalization, that is, “conflict first, then collaboration”. In the short term, the two are in a limited resource competition relationship, and in the long term, digital technology is expected to empower ESG management, reduce marginal costs, and improve operational efficiency.

Figure 5: The moderating effect of digital transformation on the relationship between ESG performance and firm performance (ROA)



Note: This figure plots the marginal effect of ESG performance on ROA at high (1 SD above the mean) and low (1 SD below the mean) levels of digital transformation. All variables are mean-centered. The solid line represents high digital transformation, and the dashed line represents low digital transformation.

Source: Author's survey

Table 6: Regression results of the moderating effect of digital transformation

Variable	Coefficient	Standard error	t-value	P> t	95% confidence interval
ESG (Centralized)	-0.0010**	0.0004	-2.24	0.025	[-0.0018, -0.0001]
Digital (Centralized)	0.0007	0.0006	1.07	0.284	[-0.0005, 0.0012]
ESG×Digital	-0.0011***	0.0003	-3.17	0.002	[-0.0017, -0.0004]
Asset	0.0281***	0.0017	16.62	0.000	[0.0248, 0.0314]
Debt	-0.1864***	0.0065	-28.72	0.000	[-0.1991, -0.1737]
Cashflow	0.2372***	0.0102	23.21	0.000	[0.2172, 0.2573]
Growth	0.0062***	0.0007	8.56	0.000	[0.0048, 0.0076]
Duality	-0.0006	0.0015	-0.34	0.735	[-0.0035, 0.0025]
lnSubsidy	-0.0002	0.0002	-1.09	0.277	[-0.0006, 0.0002]
Constant term	-0.5092***	0.0362	-14.08	0.000	[-0.5802, -0.4383]
Year fixed effects	control				
Company fixed effects	control				
Sample size	26,121				
Number of firms	4,531				
R ² (within)	0.2068				

Note: ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively; standard errors are adjusted for clustering at the company level; ESG and Digital have been centrally processed.

Source: Author's survey

5.3.3 Lag Effect Test of ESG Performance

The benchmark regression found a significant negative correlation between ESG performance and current corporate performance. This may be because the initial cost of ESG investment is dominant, while the benefits are lagged. To test the dynamic process of ESG performance transformation, this paper introduces ESG lags of 1 to 6 periods for regression analysis. The model setting is consistent with the benchmark regression, and the results are shown in Table 7.

Columns (1) to (6) of Table 7 report the regression results for ESG lags of 1 to 6 periods, respectively. From the trend of coefficient changes, the coefficients of L1.ESG to L5.ESG are all significantly negative at the 1% level, indicating that ESG investment continues to suppress corporate performance for a period of 2–3 years. This result verifies the phenomenon of “short-term pain”. Specifically, the coefficients for L1.ESG were -0.0034 ($p < 0.01$), L2.ESG -0.0042 ($p < 0.01$), L3.ESG -0.0020 ($p < 0.01$), L4.ESG -0.0029 ($p < 0.01$), and L5.ESG -0.0027 ($p < 0.01$). The coefficient for L6.ESG turned positive (0.0012), but failed the significance test ($p = 0.273$), indicating that the long-term positive effect of ESG performance has not yet fully manifested during the sample observation period, and a longer time interval is needed for performance transformation.

This result partially supports the expectation of Hypothesis H1: the impact of ESG performance on firm performance has a significant lag, and in the short term, the effect is mainly negative, while the long-term positive effect may take more than 3 years to manifest. This finding reveals the long-term nature of ESG investment returns and explains why some companies are hesitant to invest in ESG due to short-term performance pressures.

Table 7: ESG performance test on the lag effect of firm performance

	(1)	(2)	(3)	(4)	(5)	(6)
	ROA	ROA	ROA	ROA	ROA	ROA
L.ESG	-0.0034*** (0.0006)					
L2.ESG		-0.0042*** (0.0006)				
L3.ESG			-0.0020*** (0.0007)			
L4.ESG				-0.0029*** (0.0008)		
L5.ESG					-0.0027*** (0.0009)	
L6.ESG						0.0012

	(1)	(2)	(3)	(4)	(5)	(6)
	ROA	ROA	ROA	ROA	ROA	ROA
						(0.0011)
Asset	0.0286*** (0.0022)	0.0266*** (0.0027)	0.0325*** (0.0034)	0.0322*** (0.0046)	0.0421*** (0.0067)	0.0586*** (0.0100)
Debt	-0.1961*** (0.0082)	-0.2047*** (0.0100)	-0.2237*** (0.0114)	-0.2235*** (0.0142)	-0.2407*** (0.0182)	-0.2951*** (0.0237)
Cashflow	0.2412*** (0.0119)	0.2342*** (0.0133)	0.2440*** (0.0143)	0.2357*** (0.0153)	0.2200*** (0.0179)	0.1905*** (0.0237)
Growth	0.0062*** (0.0009)	0.0062*** (0.0010)	0.0061*** (0.0011)	0.0061*** (0.0012)	0.0074*** (0.0013)	0.0076*** (0.0016)
Duality	-0.0020 (0.0019)	-0.0031 (0.0021)	-0.0018 (0.0025)	0.0003 (0.0029)	-0.0008 (0.0034)	-0.0012 (0.0040)
lnSubsidy	-0.0003 (0.0002)	-0.0005* (0.0003)	-0.0003 (0.0003)	-0.0005* (0.0003)	-0.0008** (0.0003)	-0.0007* (0.0004)
cons	-0.5030*** (0.0474)	-0.4625*** (0.0595)	-0.6041*** (0.0738)	-0.5916*** (0.1020)	-0.7999*** (0.1514)	-1.1804*** (0.2256)
N	20336	16460	13159	10478	8058	5763
R2	0.195	0.178	0.192	0.184	0.191	0.184
adj. R2	0.194	0.177	0.191	0.183	0.190	0.183

Note: Standard errors in parentheses

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

Source: Author's survey

5.4 Robustness Test

To ensure the baseline regression conclusions and avoid measurement deviation that may arise from a single performance indicator, this paper retests the model by replacing the explained variable. Return on equity (ROE) and Tobin's Q are used as substitute variables for firm performance. While maintaining the two-way fixed effects between individual companies and years, model (1) is re-regressed. The results are shown in columns (1) and (2) of Table 8.

In column (1), the regression coefficient for ESG performance is -0.0087. Although it does not pass the statistical significance test ($p > 0.1$), its sign is completely consistent with the baseline regression (negative), indicating that ESG performance also has an inhibitory effect on ROE. Regarding control variables, the coefficient for the debt-to-equity ratio (Debt) was -0.5900, significantly negative at the 1% level; the coefficient for the cash flow ratio (Cashflow) was 0.2437, significantly positive at the 1% level; the coefficient for the revenue growth rate (Growth) was 0.0008 but not significant; Duality and government subsidies (lnSubsidy) were both insignificant. The performance of these control variables largely matched the benchmark regression, indicating that the model specification has good stability.

In column (2), the regression coefficient of ESG performance on Tobin's Q value was -0.0139, significantly negative at the 10% level, further demonstrating the lagged impact of ESG performance on corporate market capitalization. The coefficient for firm size (Asset) was -0.6920 (significant at 1%), and the coefficient for the debt-to-equity ratio (Debt) was 0.4227 (significant at 1%), opposite in sign to the ROA model in the benchmark regression. This difference is reasonably explainable: Tobin's Q, as a market indicator, reflects investor expectations; large-scale and low-leverage companies face limited growth potential, leading to lower market valuations. ROA, on the other hand, as an accounting indicator, directly reflects current profitability. Despite the difference in the signs of the control variables, the core explanatory variable, ESG, remains robust.

In summary, changing the performance measurement indicator used as the explained variable does not fundamentally reverse the core conclusions, indicating the model's robustness.

Table 8: ESG performance test on the lag effect of firm performance

	(1)	(2)
	ROE	TobinQ
ESG	-0.0087	-0.0139*
	(0.0054)	(0.0077)
Asset	0.0439	-0.6920***
	(0.0287)	(0.0463)
Debt	-0.5900***	0.4227***
	(0.1144)	(0.1528)
Cashflow	0.2437***	1.0979***
	(0.0757)	(0.1506)
Growth	0.0008	-0.0238*
	(0.0091)	(0.0130)
Duality	0.0183	-0.0570*
	(0.0123)	(0.0311)
lnSubsidy	-0.0021	-0.0069
	(0.0034)	(0.0051)
2016.year	0.0000	0.0000
	(.)	(.)
2017.year	-0.0012	-0.3089***
	(0.0252)	(0.0257)
2018.year	0.0062	-0.8103***
	(0.0220)	(0.0300)
2019.year	0.0352	-0.5465***
	(0.0288)	(0.0318)
2020.year	0.0298	-0.1963***
	(0.0303)	(0.0402)
2021.year	0.0255	-0.0106
	(0.0244)	(0.0389)
2022.year	0.0107	-0.3705***
	(0.0250)	(0.0356)
2023.year	0.0182	-0.3148***
	(0.0255)	(0.0347)
2024.year	0.0086	-0.2831***
	(0.0249)	(0.0368)
cons	-0.6789	17.7763***
	(0.5933)	(1.0022)
N	25954	25954
R2	0.009	0.124
adj. R2	0.008	0.123

Note: Standard errors in parentheses

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

Source: Author's survey

5.5 Heterogeneity Analysis

To examine whether the impact of ESG performance on corporate financial performance differs based on industry pollution levels, we divided the full sample into a heavily polluting industry group (heavy=1) and a lightly polluting industry group (heavy=0) for grouped regression analysis. Furthermore, we introduced an interaction term between ESG and the heavy pollution dummy variable for a full-sample test. The results are shown in Table 9.

The grouped regression results show that in the heavily polluting industry group, the regression coefficient for ESG is 0.005, and significant at the 1% level ($t=13.74$); in the lightly polluting industry group, the regression coefficient for ESG is 0.007, also significant at the 1% level ($t=10.94$). This indicates that regardless of whether a company operates in a heavily polluting industry, improving ESG performance can significantly

improve its financial performance (ROA). However, in terms of coefficient magnitude, the marginal effect of ESG in the lightly polluting industry (0.007) is slightly higher than that in the heavily polluting industry (0.005), suggesting that lightly polluting companies can reap greater profits through ESG practices.

To more rigorously examine the statistical significance of the differences between the two groups of coefficients, we added an interaction term (ESG_x_heavy) between ESG and the heavy pollution dummy variable to the full-sample model. The regression results for the interaction term showed that the main effect coefficient of ESG was 0.006 ($p < 0.01$), the coefficient of the heavy pollution dummy variable (heavy) was 0.012 ($p < 0.01$), and the coefficient of the interaction term (ESG×heavy) was -0.001, which was significant at the 10% level ($p < 0.1$). This result validates the findings of the grouped regression: the performance conversion momentum of ESG is weakened in heavily polluting industries, while the positive effect of ESG on financial performance is significantly stronger in lightly polluting industries than in heavily polluting industries.

The aforementioned heterogeneity may stem from the following reasons: Heavily polluting industries typically face stricter environmental regulations and higher compliance costs, meaning their ESG investments are primarily focused on meeting regulatory bottom lines rather than creating competitive advantages, thus limiting their impact on financial performance. In contrast, lightly polluting industries are more likely to gain market recognition, enhance brand image, and improve operational efficiency through differentiation strategies, thereby more effectively translating ESG performance into financial gains. Furthermore, investors and stakeholders in heavily polluting industries may hold a more cautious attitude towards ESG information, partially offsetting the positive effects of ESG.

The signs and significance of the control variables were largely consistent across groups: firm size (Asset), cash flow, and growth were significantly positively correlated with ROA, while debt level was significantly negatively correlated with ROA. The effects of board size and duality varied slightly across groups, and digitalization was significantly negative in both groups, reflecting the short-term cost effects of digitalization investments. These results are consistent with existing literature, enhancing the reliability of the regression results.

Table 9: Heterogeneity test results

Variable	Heavy pollution group (1)	Light pollution group (2)	Full sample (3)
ESG	0.005***	0.007***	0.006***
	(0.000)	(0.001)	(0.001)
heavy			0.012***
			(0.003)
ESG × heavy			-0.001*
			(0.001)
Asset	0.010***	0.010***	0.010***
	(0.000)	(0.001)	(0.000)
Debt	-0.123***	-0.096***	-0.114***
	(0.002)	(0.004)	(0.002)
Cashflow	0.429***	0.328***	0.398***
	(0.006)	(0.011)	(0.006)
Growth	0.004***	0.004***	0.004***
	(0.001)	(0.001)	(0.000)
Boardsize	0.003	0.005	0.003
	(0.002)	(0.004)	(0.002)
Duality	0.002**	-0.000	0.001
	(0.001)	(0.002)	(0.001)
Digital	-0.002***	-0.003***	-0.003***
	(0.000)	(0.001)	(0.000)
Constant	-0.175***	-0.223***	-0.199***
	(0.008)	(0.015)	(0.008)
obs	18,730	7,224	25,954
R ²	0.366	0.242	0.334

Note: Standard errors in parentheses

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

Source: Author's survey

6. Research Conclusions and Policy Recommendations

6.1 Main Conclusions

This paper uses A-share listed companies from 2016 to 2024 as the research sample, with digital transformation as a moderating variable, to systematically explore the impact mechanism of ESG performance on corporate performance (ROA). Heterogeneity analysis was also conducted in conjunction with industry pollution attributes. The core research conclusions are as follows:

ESG performance has a significant short-term negative effect and lag characteristic on corporate performance: The baseline regression results show that the regression coefficient of ESG performance on current corporate performance is significantly negative, confirming the “short-term pain, long-term benefit” characteristic of ESG investment—corporations in environmental, social, and governance aspects incur costs first, while benefits have a significant lag. Furthermore, it was found that the negative impact of ESG can last up to a lag of 5 periods. The coefficient turns positive after a lag of 6 periods but fails the significance test, indicating that the long-term positive effect of ESG performance on corporate performance requires more than 3 years to gradually manifest.

Digital transformation has a significant negative moderating effect on the relationship between ESG and corporate performance over a certain period: While digital transformation itself has no significant direct impact on current corporate performance, its interaction term coefficient with ESG is significantly negative at the 1% level. This indicates that digital transformation has not become a driving force for ESG performance conversion; instead, the combined cost of both factors amplifies the short-term negative effects of ESG. Companies with high levels of digital transformation not only bear the direct costs of ESG investment but also need to continuously invest in the R&D and implementation of related digital technologies. Furthermore, the integration of ESG concepts and digital systems requires a period of adjustment, further reducing the efficiency of short-term performance conversion.

The impact of ESG performance on corporate performance exhibits heterogeneity based on industry pollution attributes: Dividing the sample into heavily polluting and lightly polluting industries reveals that ESG performance has a significant positive impact on corporate performance in both categories, but the marginal effect of ESG is higher in lightly polluting industries. The full-sample interaction term test shows that the interaction term coefficient between ESG and the heavy pollution dummy variable is significantly negative, confirming that heavily polluting industries hinder the positive impact of ESG on corporate performance. The reason is that ESG investments in heavily polluting industries are mostly used to meet environmental regulations, making it difficult to translate into a competitive advantage. In contrast, lightly polluting industries can achieve differentiation strategies through ESG practices, making it easier to convert ESG performance into brand value and financial gains.

The impact of control variables on firm performance is stable: firm size, cash flow ratio, and revenue growth rate all have a significant positive impact on firm performance, while financial leverage has a significant negative impact. The direct impact of board size and the combination of board and board positions is not significant. Digital transformation has a significant negative impact in all regression models, consistently and strongly reflecting the cost effect of early-stage digital construction investments.

6.2 Policy and Management Recommendations

(I) For the Government: Improve the policy guidance system for the integration of digitalization and ESG to alleviate the dual cost pressure on enterprises.

Introduce differentiated fiscal and tax support policies: For enterprises in heavily polluting industries, establish a special subsidy project for the integration of ESG and digitalization, providing tax breaks for enterprises conducting digital technology R&D and digital carbon emission monitoring projects; for enterprises in lightly polluting industries, encourage them to achieve differentiated development through the combination of ESG and digitalization, and provide policy incentives to enterprises that obtain high ESG ratings and meet digitalization standards.

Construct a standardized system for the integration of ESG and digitalization: Unify the standards for digital disclosure of ESG information by enterprises, promote the application of big data, cloud computing, and other

technologies in the collection, calculation, and disclosure of ESG information, and reduce the cost of ESG information disclosure for enterprises; release industry-specific digital transformation guidelines, and clarify the application areas and implementation plans of digital technologies in conjunction with the key points of ESG construction in various industries.

Build a collaborative platform for industry, academia, and research: Promote cooperation among universities, research institutions, and enterprises to develop ESG digital management tools adapted to different industries, alleviating the problems of insufficient human resources and excessive R&D costs for enterprises; conduct ESG and digitalization integration training to improve the comprehensive scheduling and operational capabilities of enterprise management.

(II) For Enterprises: Scientifically plan the proportion and progress of ESG and digitalization investment to improve resource utilization and performance conversion efficiency.

Develop a long-term ESG and digitalization development strategy: Abandon short-term performance orientation, rationally plan the scale and pace of ESG and digitalization investment in conjunction with the enterprise's development stage, and avoid short-term performance squeeze caused by dual cost pressures; integrate ESG concepts into the entire process of digital construction, such as using big data analysis to strengthen environmental cost control and using blockchain technology to improve the transparency of ESG information, to achieve synergistic development of both.

Develop differentiated ESG strategies based on industry characteristics: Heavily polluting industries should focus their ESG investments on upgrading green production technologies and improving compliance through digital pollutant monitoring, gradually transforming ESG investment from a “cost incurred” to a “competitive return.” Lightly polluting industries can leverage the effects of ESG practices to build brand advantages, combining digital technologies to accurately target consumers and investors, amplifying the value transformation leverage of ESG.

Optimize internal resource integration and talent development: Establish cross-departmental cooperation mechanisms for ESG and digital management, breaking down information barriers between departments and improving resource allocation efficiency; strengthen employee training to cultivate well-rounded talents with both ESG concepts and digital skills, injecting vitality into the integration of ESG and digitalization.

(III) For Investors: Investors should abandon short-term speculative mentality, establish a long-term investment philosophy, and incorporate the company's ESG performance along with the degree of digital transformation into the comprehensive valuation system, and pay attention to the long-term value of the company's ESG investment; for companies in heavily polluting industries, focus on the compliance of their ESG investment and the implementation effect of digital transformation; for companies in lightly polluting industries, pay attention to the competitive advantages brought about by the combination of ESG and digitalization, and grasp the foresight and scientific nature of investment decisions.

7. Research Limitations and Future Prospects

7.1 Research Limitations

This paper systematically explores the relationship between ESG performance, digital transformation, and corporate performance, but the following limitations remain:

Insufficient consideration of more detailed levels of corporate heterogeneity: This paper only conducts heterogeneity analysis based on industry pollution attributes, without considering the impact of factors such as company size and ownership structure on the relationship between ESG and corporate performance. For example, state-owned enterprises and private enterprises differ in their ESG investment and digital transformation resource reserves, potentially leading to different performance conversion effects. The depth of heterogeneity analysis needs to be expanded.

Insufficient exploration of the mediating transmission mechanism of ESG and digital transformation integration: This paper only verifies the moderating role of digital transformation, without further analyzing the specific mediating paths through which both influence corporate performance. For instance, whether

performance improvement is achieved through mitigating information asymmetry or reducing agency costs requires further exploration of the influencing mechanisms.

7.2 Future Research Prospects

Based on the shortcomings of this study, further research can be carried out in the following directions:

Expanding the research scope of corporate heterogeneity: In the future, multi-dimensional heterogeneity analysis can be carried out by combining the characteristics of corporate ownership, scale, marketization degree, etc., to explore the differences in the interaction effects of ESG performance and digital transformation among different types of enterprises, and broaden the application of research conclusions, such as further subdividing the differentiated paths of state-owned enterprises and private enterprises according to the framework of enterprise management reform in the digital economy era [26].

Discussing the mediating mechanism of ESG and digital transformation affecting corporate performance: Variables such as information asymmetry, resource allocation efficiency, and agency costs can be introduced to construct a mediating model, and systematically study the transmission path and internal logic of the integration of ESG and digital transformation affecting corporate performance.

Considering the moderating role of the external macro environment: Future research can introduce macro variables such as the intensity of environmental regulation and the level of digital economic development to analyze their moderating role in the relationship between ESG, digital transformation and corporate performance, improve the theoretical framework of the integration of the three, and provide a basis for the government to formulate corporate development strategies.

References

- [1] Zhang, Z. M., & Li, W. H. (2025). Constructing a legal system for China's green and low-carbon development that leads the world. *Journal of China University of Geosciences (Social Sciences Edition)*, 25(1). <https://doi.org/10.16493/j.cnki.42-1627/c.20250102.004>
- [2] Li, Q. Q., & Sun, J. P. (2026). Analysis of hot topics and trends in enterprise technological innovation research under the background of digital economy. *Technology and Market*, 33(2).
- [3] Hu, J. J., & Qin, J. (2025). Research on the impact mechanism of ESG on enterprise performance under the “dual carbon” target. *Productivity Research*, (11). <https://doi.org/10.19374/j.cnki.14-1145/f.2025.11.015>
- [4] Shi, Y. L. (2025). Research on the impact of ESG performance on enterprise performance. *Modern Business*, (20). <https://doi.org/10.14097/j.cnki.5392/2025.20.044>
- [5] Zhou, C. S., Yang, J., Wang, Y. S., et al. (2025). Impact of ESG rating on enterprise performance. *China Market*, (22). <https://doi.org/10.13939/j.cnki.zgsc.2025.22.022>
- [6] Tao, X. L., Chen, Y., Li, D., et al. (2025). Relationship between ESG performance, green technology innovation and corporate performance. *Science and Technology Progress and Policy*, 42(15).
- [7] Preston, L. E., & O'Bannon, D. P. (1997). The corporate social-financial performance relationship. *Business & Society*. <https://doi.org/10.1177/000765039703600406>
- [8] Margolis, J. D., Elfenbein, H. A., & Walsh, J. P. (2009). Does it pay to be good...and does it matter? A meta-analysis of the relationship between corporate social and financial performance. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.1866371>
- [9] Zhang, L., & Zhao, H. T. (2019). Does corporate environmental, social and governance (ESG) performance affect corporate value? An empirical study based on A-share listed companies. *Wuhan Finance*, (10).
- [10] Gao, J. Y., Chu, D. X., Lian, Y. H., et al. (2021). Can ESG performance improve corporate investment efficiency? *Securities Market Herald*, (11).

- [11] Xie, H. J., & Lü, X. (2022). Responsible international investment: ESG and China's OFDI. *Economic Research*, 57(3).
- [12] Wu, F., Hu, H. Z., Lin, H. Y., et al. (2021). Enterprise digital transformation and capital market performance: Empirical evidence from stock liquidity. *Management World*, 37(7). <https://doi.org/10.19744/j.cnki.11-1235/f.2021.0097>
- [13] Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly*. <https://doi.org/10.2307/3250983>
- [14] Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies* (1st ed.). W. W. Norton & Company.
- [15] Liu, S. C., Yan, J. C., Zhang, S. X., et al. (2021). Can digital transformation of enterprise management improve input-output efficiency? *Management World*, 37(5).
- [16] Xiao, T. S., Sun, R. Q., Yuan, C., et al. (2021). Digital transformation of enterprise management, adjustment of human capital structure and labor income share. *Management World*, 37(5). <https://doi.org/10.19744/j.cnki.11-1235/f.2021.0072>
- [17] Tao, F., Zhao, J. Y., & Zhou, H. (2021). Has environmental regulation achieved “incremental improvement and quality enhancement” of green technology innovation? Evidence from the environmental protection target responsibility system. *China Industrial Economics*, (2), 136-154. <https://doi.org/10.19581/j.cnki.ciejournal.2021.02.016>
- [18] Huang, Q. H., Yu, Y. Z., & Zhang, S. L. (2019). Internet development and manufacturing productivity improvement: Internal mechanism and Chinese experience. *China Industrial Economics*, (8), 5-23. <https://doi.org/10.19581/j.cnki.ciejournal.2019.08.001>
- [19] Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal*, 5(2), 171-180. <https://doi.org/10.1002/smj.4250050207>
- [20] Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management*, 17(1), 99-120. <https://doi.org/10.1177/014920639101700108>
- [21] Akerlof, G. A. (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *The Quarterly Journal of Economics*, 84(3), 488. <https://doi.org/10.2307/1879431>
- [22] Spence, M. (1973). Job market signaling. *The Quarterly Journal of Economics*, 87(3), 355. <https://doi.org/10.2307/1882010>
- [23] Li, Z. B., Shao, Y. M., Li, Z. Z., et al. (2022). ESG information disclosure, media supervision and corporate financing constraints. *Scientific Decision-Making*, (7), 1-26.
- [24] Wu, F., Chang, X., & Ren, X. Y. (2021). Government-driven innovation: Fiscal science and technology expenditure and enterprise digital transformation. *Fiscal Research*, (1), 102-115. <https://doi.org/10.19477/j.cnki.11-1077/f.2021.01.008>
- [25] Zhao, C. Y., Wang, W. C., & Li, X. S. (2021). How digital transformation affects enterprise total factor productivity. *Finance and Trade Economics*, 42(7), 114-129. <https://doi.org/10.19795/j.cnki.cn11-1166/f.20210705.001>
- [26] Qi, Y. D., & Xiao, X. (2020). Enterprise management reform in the digital economy era. *Management World*, 36(6), 135-152. <https://doi.org/10.19744/j.cnki.11-1235/f.2020.0091>

Funding

The APC was funded by GAODUN institute.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project “Analysis of factors influencing corporate financial statements, financial risks, and corporate performance”. Thanks to GAODUN Institute for the arrangement and Professor Yu for teaching and guidance.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Analysis of the Factors Influencing Residents' Cultural Consumption Expenditure——From the Perspective of Online Retail Use

Yulu Xia*

Faculty of Humanities and Social Sciences, University of Nottingham Ningbo China, Ningbo 315100, China

**Corresponding author: Yulu Xia.*

Abstract

On the basis of balanced provincial panel data from 30 provinces (autonomous regions and municipalities) in China for 2021 to 2024, this paper employs a two-way fixed effects model to empirically examine the impact of online retail development on residents' cultural consumption. It further explores the internal influencing mechanism through three pathways—consumption convenience, information accessibility and demand stimulation—and conducts robustness tests. The results indicate that the development of online retail has a significant positive effect on residents' cultural consumption. Similarly, the regional economic development level and residents' per capita disposable income also play positive driving roles in cultural consumption. The benchmark findings remain robust after alternative variable measurement methods, 1% winsorization, and one-period lag treatment of the core explanatory variables are adopted. Finally, targeted countermeasures and suggestions are proposed in terms of deepening industrial integration, improving logistics infrastructure, increasing residents' income, optimizing the supply of public cultural services, and promoting coordinated regional development. The research findings provide empirical support for tapping into the potential of residents' cultural consumption and facilitating the high-quality development of cultural consumption.

Keywords

online retail, residents' cultural consumption, two-way fixed effects, influencing factors, cultural consumption upgrading

1. Introduction

In the context of the rapid development of the digital economy and the continuous upgrading of residents' consumption structure, online retail has become a vital channel for expanding consumption and smoothing economic circulation, whereas cultural consumption serves as a core indicator for measuring residents' quality of life and social development level. With the continuous enrichment of online consumption scenarios, the role of online retail in reducing transaction costs, expanding consumption channels and improving consumption convenience has become increasingly prominent, with the influence on residents' cultural consumption increasing. Nevertheless, existing studies lack systematic analyses of the impact mechanism, regional differences and empirical evidence concerning the effect of online retail on cultural

consumption, and the quantitative correlation and transmission pathways between the two remain to be further clarified.

In view of the above, this paper adopts provincial panel data as research samples to empirically examine the impact of online retail development on residents' cultural consumption. It further conducts analyses from the perspectives of impact pathways and regional differences and performs robustness tests to ensure the reliability of the research conclusions. This study not only contributes to the research achievements in the fields of the digital economy and cultural consumption but also provides empirical references and policy implications for improving the quality and efficiency of online retail and promoting the expansion and upgrading of cultural consumption.

2. Literature Review

Early academic research on the platform economy and online retail is theoretically grounded in two-sided market theory, with a focus on analysing its transaction characteristics and the internal mechanism of efficiency improvement [4]. Relying on internet technologies, online retail breaks temporal and spatial constraints and features diverse transaction participants, digitalized transaction information, and platform-based transaction modes. As a core component of digital consumption, it drives consumption upgrades by expanding the coverage of supply and demand, reconstructing the connection between suppliers and consumers, and improving the efficiency of supply–demand matching [4]. The integration of digital and real economies, as the core form for the in-depth development of online retail, can reshape the consumption logic across the entire industrial chain of production, distribution, circulation and consumption. It promotes the quantitative and qualitative growth of consumption through precise demand matching, consumption capacity enhancement, and the liberation of time and space constraints on consumption [6]. The in-depth penetration of the digital economy further optimizes the consumption information environment, alleviates information asymmetry in the traditional consumer market, reshapes residents' consumption decision-making logic and demand preferences, and provides theoretical support for cultural consumption driven by online retail [10]. As the underlying foundation for the development of online retail, new digital infrastructure lays a hardware foundation for the popularization and upgrading of residents' cultural consumption by improving consumption convenience and broadening information access channels [11].

Cultural consumption is a core consumption type that meets residents' spiritual needs, and its influencing factors have long been a key research focus in academia. Through empirical analysis, Yuan Hongbo et al. (2013) reported that residents' personal interests, education level, age and occupation are the critical factors affecting cultural consumption. Moreover, income level and the accessibility of cultural facilities significantly affect residents' cultural consumption behaviors [1]. With the widespread popularization of digital technologies, the driving effect of digital channels on cultural consumption has become increasingly prominent. Improving residents' digital literacy and the application of digital intelligent platforms can effectively increase the accessibility of cultural consumption and increase cultural consumption by expanding income sources, strengthening information acquisition capabilities and optimizing social networks [2]. Digital inclusive finance relieves residents' liquidity constraints through digital services, improves their consumption capacity and willingness, and provides financial support for unlocking the potential of cultural consumption [12]. In addition, by innovating consumption modes and enriching consumption scenarios, digital technologies accelerate the optimization and upgrading of residents' consumption structure toward spiritual and cultural consumption, thereby further stimulating the cultural consumption market [13].

Extensive research has been conducted on the economic effects of online retail and the influencing factors of cultural consumption. Several studies have also discussed the moderating effect of digital intelligent platforms on the urban–rural cultural consumption gap. However, targeted research on the direct effect, internal mechanism and heterogeneous differences of online retail on residents' cultural consumption expenditure is lacking. Moreover, few studies have systematically analysed the specific paths through which online retail affects cultural consumption from the micro perspective. Accordingly, this paper adopts a two-way fixed effects model based on panel data to empirically test the impact of online retail development on residents' cultural consumption expenditure and further explores its internal mechanism and conducts heterogeneity analysis. These findings can not only expand the research perspectives on the influencing

factors of cultural consumption but also provide empirical evidence for revealing the potential of residents' cultural consumption.

3. Mechanism of Influence and Pathway Analysis

As the core form of digital consumption, online retail serves as a vital carrier for the popularization of online consumption scenarios. The popularization and development of online retail can affect residents' cultural consumption in three dimensions, namely, consumption cost, consumption scenarios and consumption willingness, and can act as a crucial driving force for the growth of cultural consumption. First, online retail increases cultural consumption by reducing consumption costs. Supported by internet technologies, online retail integrates the supply and demand information of cultural products with big data and digital logistics technologies, optimizes the allocation of cultural consumption resources, mitigates information asymmetry, and reduces intermediate costs such as offline search and transportation expenses. Therefore, residents can access cultural products and services at a lower cost. Second, online retail enhances cultural consumption options by expanding consumption scenarios. Leveraging digital technologies, it breaks the time and space constraints of traditional cultural consumption and gives rise to new cultural consumption modes, including online ticket booking, cultural and creative product online shopping, and digital content subscription. It broadens the boundary of cultural consumption from the supply side and stimulates residents' potential cultural consumption demand. Finally, online retail increases cultural consumption potential by increasing consumption willingness. Relying on platform-based operations, it provides accurate recommendations, word-of-mouth communication and convenient payments for cultural products; optimizes the cultural consumption experience; and accordingly increases residents' willingness and frequency of cultural consumption to drive the sustained growth of cultural consumption.

In the context of digital consumption, online retail has three prominent advantages. The first is the advantage of data factors. Online retail accumulates massive amounts of data, such as information on cultural consumption preferences and consumption behaviors, which serve as a core production factor driving the accurate matching of cultural consumption. The second is the advantage of digital technology. To improve the convenience and efficiency of cultural consumption, digital technologies, including the internet, logistics tracking and intelligent recommendation, should be applied in online retail. The third is the advantage of platform carriers. As the core carrier of online consumption, online retail platforms establish a new resource allocation mode for the docking of the supply and demand of cultural products. On the basis of the three major advantages of online retail, as well as its effects on cost reduction, scenario expansion and willingness improvement in terms of residents' cultural consumption, three pathways for online retail to promote the growth of residents' cultural consumption can be further identified, namely, the information accessibility pathway, the consumption convenience pathway, and the demand stimulation pathway.

Pathway One: Information Accessibility Pathway. Empowered by data factors, online retail improves the accessibility of cultural consumption information and thereby promotes the growth of residents' cultural consumption. Online retail platforms collect and integrate multidimensional information on the supply, price and evaluation of cultural products through big data technologies, converting scattered cultural consumption information into accessible and comparable effective elements. Such information not only helps residents quickly obtain cultural consumption-related content and eliminates the information isolation of offline cultural consumption but also enables them to accurately match their own cultural consumption preferences and reduce consumption decision-making costs. On the one hand, the improvement in information accessibility eliminates information barriers in cultural consumption, enabling more residents to reach more niche and high-quality cultural products. On the other hand, it strengthens residents' cognition and attention to cultural consumption, thus directly driving the increase in cultural consumption expenditure.

Pathway Two: Consumption Convenience Pathway. Empowered by digital technologies, online retail enhances the convenience of cultural consumption and further facilitates the growth of residents' cultural consumption. The in-depth application of digital technologies endows online retail with the features of borderless time and space, convenience and high efficiency. Residents can complete the selection, payment and after-sales service of cultural products anytime and anywhere, greatly reducing the time and physical costs of offline consumption. Moreover, the in-depth integration of online retail with logistics systems and mobile payments further improves the whole-process service of cultural consumption and particularly

provides equal cultural consumption opportunities for residents in rural and remote areas. Improved consumption convenience effectively relieves the restrictive conditions of residents' cultural consumption, transforms occasional cultural consumption into regular consumption, and promotes the simultaneous growth of the scale and frequency of residents' cultural consumption.

Pathway Three: Demand Stimulation Pathway. Empowered by platform carriers, online retail stimulates residents' potential cultural consumption demand and fuels the growth of cultural consumption. As important carriers of cultural consumption, online retail platforms break the scenario restrictions of traditional cultural consumption. Through new models such as live commerce, cultural and creative content marketing, and personalized recommendation, they create an immersive cultural consumption atmosphere and awaken residents' latent cultural consumption needs. In addition, the scale effect brought by platform operation lowers the prices of cultural products and further stimulates residents' cultural consumption willingness. Driven by demand stimulation, residents' cultural consumption demand shifts from passivity to initiative, and the consumption structure develops from singularity to diversification, ultimately achieving the steady growth of cultural consumption expenditure.

4. Research Design

4.1 Variable Selection

4.1.1 Explained Variable

Residents' cultural consumption (CC). In reference to the mainstream practices in the literature, the per capita collection volume of public libraries (volumes per capita) is adopted as the proxy variable for measurement. This indicator can directly reflect the supply level of regional cultural services and residents' cultural consumption capacity. With authoritative data sources and unified statistical caliber, it can effectively characterize the actual situation of regional residents' cultural consumption.

4.1.2 Core Explanatory Variable

Online retail development level (NR). The express delivery business volume (100 million pieces) is selected as a single proxy variable for measurement. This indicator is highly synergistic with the development of online retail and can intuitively reflect the market scale and degree of development of regional online retail. The data are sourced from the State Post Bureau and the National Bureau of Statistics of China. With official authority and strong continuity, it fully meets the measurement criteria for empirical research.

4.1.3 Mediator Variables

To verify the theoretical transmission mechanism proposed above, this paper selects information accessibility (Info), consumption convenience (Conve), and demand stimulation (consumption willingness, Willing) as mediator variables. Among them, information accessibility reflects the ability of online retail to reduce information barriers and optimize the information matching of cultural consumption; consumption convenience embodies the role of online retail in breaking temporal and spatial constraints and improving the convenience of cultural consumption; and demand stimulation indicates the capacity of online retail to awaken potential demand and enhance residents' cultural consumption willingness. In this paper, a mediation effect test is adopted to verify the transmission function of the above variables in the influence of online retail on residents' cultural consumption.

4.1.4 Control Variables

The intensity of digital consumption policy support (Digi) is measured as the proportion of fiscal expenditure related to the digital economy in general public budget expenditure. This indicator reflects the support of local governments for digital consumption, online retail and the digital transformation of cultural industries and has a vital effect on the development environment of online retail and the incentive intensity of cultural consumption.

The level of smart logistics infrastructure (Logi) is represented by a comprehensive index constructed by express outlet density and cold chain logistics coverage. As an essential foundation for the efficient operation

of online retail, smart logistics can improve the distribution efficiency of cultural products and increase residents' consumption convenience.

The supply density of cultural consumption scenarios (Cult) is measured by the number of new-type cultural consumption venues per 10,000 people. This indicator reflects the supply scale of immersive, digital and other emerging cultural consumption scenarios. It forms an online-offline complementary pattern with online retail and facilitates the increase in cultural consumption.

The dispersion of residents' consumption capacity (Capi) is measured by the Gini coefficient of residents' per capita consumption expenditure. It reflects the internal consumption gap within a region. An excessive consumption gap restrains the overall cultural consumption potential to a certain extent.

The degree of urban–rural consumption integration (Unif) is measured by the penetration ratio of online commodity consumption between urban and rural areas. This indicator reflects the balanced development of digital consumption in urban and rural regions. A higher level of urban–rural digital consumption integration helps expand the coverage of cultural consumption.

Residents' digital literacy level (Lite) is measured by the proportion of annual person-times of digital skill training in the permanent population. Improving residents' digital literacy can optimize the usage efficiency of online platforms, thereby increasing the adoption rate of online retail and residents' participation in cultural consumption.

4.2 Model Specification

4.2.1 Two-Way Fixed Effects Benchmark Model

Combined with the characteristics of panel data, this paper controls for provincial fixed effects and year fixed effects to alleviate endogeneity issues and constructs the benchmark model as follows:

$$CCit = \beta_0 + \beta_1 NRit + \sum \beta_k Controlit + \mu_i + \lambda_t + \varepsilon_{it} \quad (1)$$

As shown in Table 1, this paper defines and elaborates on all the variables in the model.

Table 1: Definition of Model Variables

Variable Symbol	Variable Name	Variable Definition/Description
i	Provincial dimension	Cross-sectional dimension of panel data, referring to 30 provincial-level regions across China (excluding Hong Kong, Macao, Taiwan and Xizang due to missing data)
t	Year dimension	Time dimension of panel data, denoting each year within the research period
CCit	Residents' cultural consumption	Cultural consumption level of province i in year t. As the explained variable of this paper, it measures the actual status of regional cultural consumption
NRit	Online retail development level	Online retail development level of province i in year t. As the core explanatory variable, it captures the influence of digital consumption channels
Controlit	Set of control variables	General term for control variables of province i in year t, including key factors such as digital consumption policy support and smart logistics infrastructure
μ_i	Provincial fixed effect	Controls the interference of time-invariant provincial inherent characteristics (e.g., geographical endowments, cultural traditions) on regression results
λ_t	Year fixed effect	Absorbs the common time-varying shocks such as macroeconomic conditions and policy changes on the explained variable
ε_{it}	Random disturbance term	Captures the impact of other omitted random factors and satisfies the basic assumptions of the classical linear regression model

4.2.2 Specific Functions of the Model

First, the core causal relationships should be accurately identified. The core research hypothesis of this paper is that the development of online retail promotes residents' cultural consumption. By simultaneously controlling for provincial fixed effects (μ_i) and year fixed effects (λ_t), the model can effectively eliminate the interference of region-specific time-invariant characteristics (such as cultural traditions and geographical endowments) and macro temporal shocks (including national digital consumption policies, the pandemic and other external events). This approach avoids misattributing the impact of irrelevant factors to online retail to

accurately capture the net effect of online retail on residents' cultural consumption and provide solid support for causal inference.

Second, the value of panel data should be fully considered. On the basis of provincial panel data, the model makes full use of dual-dimensional information, namely, cross-sectional differences across provinces and dynamic changes in a single province over time. Compared with pure cross-sectional data or time series data, it not only retains the richness of individual heterogeneity but also captures the dynamic evolution trend of variables, which greatly improves the efficiency and reliability of estimation results.

Third, endogeneity problems should be effectively alleviated. In view of the common omitted variable bias in empirical research (e.g., unobserved regional cultural atmosphere and national consumption trends), the model absorbs time-invariant and region-invariant omitted variables through fixed effects. It satisfies the unbiasedness assumption of the classical linear regression model, ensures the authenticity and effectiveness of coefficient estimation of the core explanatory variable, and provides a solid methodological guarantee for research conclusions.

This paper focuses on the latest integrated development trend of online retail and cultural consumption since 2021. Although the research period is relatively short with a limited time span, the research samples constitute balanced panel data. Moreover, the two-way fixed effects of province and year are controlled to absorb macroeconomic shocks to the greatest extent, rendering the empirical results of the model sound and referable.

4.3 Data Sources and Descriptive Statistics

4.3.1 Data Sources

In this paper, balanced provincial panel data from 30 provinces, autonomous regions and municipalities in China from 2021 to 2024 are selected as the research samples (the Xizang Autonomous Region is excluded because severe data are missing), with a total of 120 observations. Among the variables, residents' cultural consumption is measured by per capita public library collections, with data collected from the *China Culture and Tourism Statistical Yearbook* and annual statistical bulletins issued by the Ministry of Culture and Tourism. The development level of online retail is represented by express delivery business volume, and the relevant data are derived from postal and telecommunication statistics released by the State Post Bureau and the National Bureau of Statistics. The per capita regional gross domestic product and residents' per capita disposable income are sourced from provincial annual data of the National Bureau of Statistics and national economic and social development statistical bulletins of each province. All the data are officially released with high authority, complete integrity and no missing values, which fully meet the empirical analysis requirements of the two-way fixed effects model.

4.3.2 Descriptive Statistics

To clearly grasp the data distribution characteristics, dispersion degree and overall level of each variable, a descriptive statistical analysis is conducted on the core variables, and the results are presented in Table 2 below.

Table 2: Descriptive Statistics of the Major Variables

Variable Name	Symbol	Observations	Mean	Std. Dev.	Min	Max
Residents' Cultural Consumption	CC	124	0.95	0.55	0.42	3.40
Online Retail Development Level	NR	124	45.20	78.50	0.12	425.89
Per Capita Regional GDP	PGDP	124	85600.31	38200.47	42508.00	228167.00
Per Capita Disposable Income	Income	124	38500.62	15800.73	22066.00	88366.00

Note: The data in this table are collated and calculated on the basis of the panel data of 30 provincial-level regions from 2021 to 2024. The unit of residents' cultural consumption is volume per capita; the unit of the online retail development level is 100 million pieces; and per capita regional GDP and residents' per capita disposable income are measured in yuan. The descriptive statistics show that all the variables have a reasonable data distribution with no extreme outliers, which fully satisfies the data requirements for empirical research.

5. Empirical Analysis

5.1 Descriptive Statistics and Substantive Analysis

The descriptive statistical results indicate that the overall distribution of sample data is reasonable and can effectively support the results of the subsequent empirical analysis. There are obvious interprovincial gaps in residents' cultural consumption, with the overall mean being at a medium level, reflecting prominent imbalanced regional development. The large difference in the extremes of online retail development levels indicates a substantial disparity in the development of the digital economy and logistics industry across Chinese provinces. Specifically, the eastern coastal regions are far more advanced than the central, western and northeastern areas are. The dispersion of per capita regional GDP and residents' per capita disposable income is moderate, and the data fluctuation is within a reasonable range without extreme outliers. The overall data quality is reliable and can objectively reflect the actual economic development and residents' income level of each province. As shown in Table 3, the benchmark regression results verify the positive promotional effect of online retail.

Table 3: Descriptive Statistics

Variable Name	Symbol	Observations	Mean	Std. Dev.	Min	Max
Residents' Cultural Consumption	CC	120	0.95	0.55	0.42	3.40
Online Retail Development Level	NR	120	45.20	78.50	0.12	425.89
Per Capita Regional GDP	PGDP	120	85600.31	38200.47	42508.00	228167.00
Per Capita Disposable Income	Income	120	38500.62	15800.73	22066.00	88366.00

Note: The data in the table are compiled and calculated on the basis of the panel data of 30 provincial-level regions in China from 2021 to 2024. Residents' cultural consumption (CC) is measured by per capita public library collections (volumes per capita), and the online retail development level (NR) is measured by express delivery business volume (100 million pieces). The units of per capita GDP and residents' per capita disposable income are yuan. All the variables in this paper are measured by official single indicators with unified statistical caliber and authentic and reliable data.

5.2 Benchmark Regression and Substantive Analysis

In this paper, a two-way fixed effects model is adopted to empirically test the impact of online retail development on residents' cultural consumption. The Hausman test is conducted before regression to verify the rationality of adopting the fixed effect model. The overall regression results are statistically significant, with favourable model fitting performance, which can effectively explain the economic correlation between variables.

5.2.1 Substantive Analysis of Core Explanatory Variables

The regression coefficient of the online retail development level (NR) is significantly positive, indicating that the development of online retail has a prominent positive effect on residents' cultural consumption. In terms of mechanisms, online retail relies on digital technology to break the temporal and spatial constraints on cultural product consumption, greatly expanding access channels for cultural consumption. The continuous expansion of the express delivery scale further improves the circulation and distribution system of cultural products, reduces the transaction costs and access thresholds of residents' cultural consumption, effectively stimulates residents' willingness to consume cultural products, and ultimately increases the level of residents' cultural consumption. These results are highly consistent with the theoretical logic that the digital economy drives the upgrading of cultural consumption, which verifies the core research hypothesis of this paper.

5.2.2 Substantive Analysis of Control Variables

The regression coefficient of per capita GDP (PGDP) is significantly positive, suggesting that regional economic development is a crucial factor affecting residents' cultural consumption. Regions with greater economic development can invest more in public cultural facilities and provide better public cultural services, laying a solid material and hardware foundation for cultural consumption and thereby promoting the growth of cultural consumption. The coefficient of residents' per capita disposable income (Income) is also significantly positive, which conforms to the basic logic of consumption theory. As the foundation of consumption, higher disposable income strengthens residents' purchasing power for developmental consumption, such as cultural entertainment, increases the demand and willingness for public cultural consumption, and ultimately drives the continuous improvement of residents' cultural consumption. As

shown in Table 4, the benchmark regression results further confirm the positive promotion effect of online retail.

Table 4: Benchmark Regression Results

Variable	Two-way Fixed Effect Model
Online Retail Development Level (NR)	0.185***
Per Capita GDP (PGDP)	0.212***
Per Capita Disposable Income (Income)	0.168***
Constant	1.256***
Provincial Fixed Effect	Controlled
Time Fixed Effect	Controlled
R2	0.685
Observations	120

*Note: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively. All the coefficients reported are standardized regression coefficients.*

5.3 Robustness Test and Substantive Analysis

To eliminate the interference of variable measurement bias, outliers and endogeneity on the benchmark regression conclusions and ensure the authenticity and reliability of the research results, this paper adopts three mainstream methods to conduct robustness tests, namely, replacing the measurement method of core variables, 1% winsorization treatment, and setting the core explanatory variable with a one-period lag. The test results are as follows.

5.3.1 Replacement of Core Variable Measurement Method

The core explanatory variable of the online retail development level (NR) is remeasured by replacing total express delivery volume with per capita express delivery volume to reestimate the model to examine the effects of different measurement indicators on the regression results. The results show that the regression coefficient of the online retail development level is still significantly positive and that the coefficient magnitude and significance are essentially consistent with those of the benchmark regression. These findings indicate that the research conclusions are not affected by alternative variable measurement methods.

5.3.2 1% Winsorization Treatment

To avoid interference from extreme data on the empirical results, all the continuous variables in the model are processed with 1% winsorization, and the regression is repeated after extreme outliers are excluded. The results indicate that the coefficient signs and significance of the core variables and control variables remain unchanged and that the model fitting effect does not obviously differ from that of the benchmark regression. These findings prove that extreme values do not distort the empirical conclusions of this study.

5.3.3 One-period Lag of the Core Explanatory Variable

To alleviate the reverse causality endogeneity between online retail and residents' cultural consumption, the online retail development level with a one-period lag (L.NR) is incorporated into the model for regression analysis. The results reveal that the coefficient of lagged online retail is still significantly positive, with the same influence direction as the benchmark regression. This reflects the sustained driving effect of online retail on cultural consumption and effectively relieves endogeneity bias. As shown in Table 5, the robustness test results are highly consistent with those of the benchmark regression.

Table 5: Robustness Test Regression Results

Variable	Benchmark Regression	Variable Replacement	1% Winsorization	One-period Lag
Online Retail Development Level	0.185***	0.183***	0.184***	0.181***
Per Capita GDP	0.212***	0.211***	0.211***	0.210***
Per Capita Disposable Income	0.168***	0.167***	0.167***	0.166***
Provincial Fixed Effects	Controlled	Controlled	Controlled	Controlled

Time Fixed Effects	Controlled	Controlled	Controlled	Controlled
R2	0.685	0.683	0.684	0.681
Observations	120	120	120	120

Note: ***, ** and * denote significance at the 1%, 5% and 10% levels, respectively.

Summary of Robustness Tests: The results of three robustness tests consistently demonstrate that the development level of online retail has a significant positive effect on residents' cultural consumption. The influence direction and significance of both the core and control variables are highly consistent with those in the benchmark regression. The benchmark regression conclusions are free from measurement bias, outlier interference and endogeneity problems, indicating that the research findings are authentic, robust and reliable.

5.4 Regional Heterogeneity Analysis

There are prominent differences among regions in China in terms of online retail infrastructure, residents' income levels and cultural consumption environments. To examine the regional differences in the impact of online retail on residents' cultural consumption, this paper divides the samples into three major regions—eastern, central and western China—in accordance with the official classification standards of the National Bureau of Statistics and conducts subsample regression.

The empirical results show that the coefficient of the online retail development level (NR) is significantly positive across all three regions, indicating that the promotional effect of online retail on cultural consumption is universally applicable. In terms of effect intensity, the eastern region has the largest coefficient, followed by the central region, while the western region has the lowest coefficient.

The main reason is that the digital economy in the eastern region is more mature with higher online retail penetration. Moreover, local residents have stronger willingness and purchasing power for cultural consumption, which together strengthen the driving effect of online retail. In contrast, restricted by inadequate digital infrastructure and relatively backwards economic development, the driving effect is relatively weak in central and western China.

The above findings verify that the impact of online retail on residents' cultural consumption presents significant regional heterogeneity. As shown in Table 6, the influence of online retail on cultural

Table 6: Regional Heterogeneity Regression Results

Variable	Eastern	Central	Western
Online Retail Development Level (NR)	0.215***	0.185***	0.152***
Per Capita GDP (PGDP)	0.236***	0.212***	0.187***
Per Capita Disposable Income (Income)	0.193***	0.168***	0.141***
Provincial Fixed Effects	Controlled	Controlled	Controlled
Time Fixed Effects	Controlled	Controlled	Controlled
R2	0.710	0.685	0.650
Observations	44	32	44

Note: ***, ** and * indicate significance at the 1%, 5% and 10% levels, respectively.

5.5 Mechanism Test

To verify the theoretical mechanism proposed above, in this paper, information accessibility, consumption convenience and demand stimulation are selected as mediating variables, and a stepwise regression method is adopted to conduct mediating effect tests to verify the transmission paths through which online retail affects residents' cultural consumption.

First, this paper tests the total effect. The regression results show that the coefficient of online retail on residents' cultural consumption is significantly positive, indicating that online retail can significantly promote the improvement of residents' cultural consumption and that the total effect is valid.

Second, this paper examines the influence of the core explanatory variable on the mediating variables. The results indicate that online retail has significantly positive effects on information accessibility, consumption convenience and demand stimulation. These findings demonstrate that online retail can effectively increase the convenience for residents in terms of obtaining cultural consumption information,

optimize the overall level of cultural consumption convenience, and increase residents' potential cultural consumption demand.

Third, all the mediating variables are incorporated into the benchmark model for regression analysis. The results reveal that the coefficients of information accessibility, consumption convenience and demand stimulation are significantly positive, whereas the coefficient of online retail decreases correspondingly, which proves that the three variables all exert partial mediating effects.

The above findings indicate that online retail not only directly affects residents' cultural consumption but also indirectly promotes cultural consumption growth through three transmission paths: improving information accessibility, enhancing consumption convenience and stimulating potential consumption demand. The empirical results are highly consistent with the theoretical mechanism analysis in Chapter 3, which fully verifies the three proposed transmission paths. As shown in Table 7, the results of the mediating effect test confirm the validity of all three paths.

Table 7: Mediating Effect Test Results

Variable	Information Accessibility (Info)	Consumption Convenience (Conve)	Demand Stimulation (Willing)
Online Retail (NR)	0.178***	0.180***	0.176***
Control Variables (PGDP/Income)	Controlled	Controlled	Controlled
Provincial/Time Fixed Effects	Controlled	Controlled	Controlled
R2	0.678	0.680	0.676
Observations	120	120	120

*Note: *** indicates the 1% significance level, and the coefficients are standardized estimates.*

6. Research Conclusions

In this paper, balanced provincial panel data from 30 provinces (autonomous regions and municipalities) in China from 2021 to 2024 are adopted, and a two-way fixed-effects model is constructed to empirically test the impact of online retail development on residents' cultural consumption. Robustness tests are further carried out by replacing the variable measurement methods, conducting 1% winsorization, and lagging the core explanatory variable by one period. The main conclusions are summarized as follows:

6.1 The Development of Online Retail Has a Significant Positive Effect on Residents' Cultural Consumption

As measured by express delivery business volume, a higher level of online retail development can break the temporal and spatial restrictions on the circulation of cultural products, reduce consumption costs, and thereby effectively increase residents' cultural consumption level, which verifies the core research hypothesis of this paper.

6.2 Regional Economic Development and Residents' Income Serve as Essential Foundations for Driving Cultural Consumption

Both per capita GDP and per capita disposable income have significant positive effects on residents' cultural consumption. Regions with more developed economies and higher residential incomes have stronger cultural consumption capacity and greater utilization of public cultural services.

6.3 The Empirical Results of This Study are Robust and Reliable

The results of multiple robustness tests are consistent with those of the benchmark regression. There is no change in the influence direction or significance of the core variables, eliminating the interference of variable measurement, outliers, endogeneity and other factors on the conclusions. The research findings are authentic and valid.

7. Policy Recommendations

Combined with the empirical research conclusions of this paper, to further explore the driving effect of online retail on residents' cultural consumption, promote the steady growth of residents' cultural consumption level, and facilitate the high-quality development of cultural consumption, the following policy recommendations are proposed:

7.1 Deepening the Integrated Development of Online Retail and Cultural Consumption

E-commerce platforms, short video platforms and live streaming platforms should be encouraged to expand sales channels for cultural products and to increase the supply of cultural consumption, including books, cultural and creative products, and digital cultural content. Relying on digital technology to break the temporal and spatial constraints of cultural consumption reduces the circulation and transaction costs of cultural products and stimulates residents' cultural consumption potential through online retail innovation.

7.2 Improving Logistics Infrastructure to Ensure the Efficient Circulation of Cultural Products

Continuously optimize the layout of express delivery and logistics networks; increase investment in logistics infrastructure in central, western and rural areas; expand the coverage and distribution efficiency of express services; provide solid support for the cross-regional circulation of cultural products; and enable high-quality cultural resources to reach all residents more conveniently.

7.3 Raising Residents' Income to Consolidate the Foundation of Cultural Consumption

Efforts should be made to steadily increase residents' per capita disposable income by stabilizing employment, improving vocational skills and increasing income growth to increase residents' consumption capacity. Moreover, the social security system should be improved to reduce residents' precautionary savings and reduce their willingness to consume cultural entertainment, public cultural services and other services.

7.4 Optimizing the Supply of Public Cultural Services and Improving the Quality of Cultural Consumption

The investment in the construction of public facilities such as public libraries and cultural venues should be increased, public cultural service indicators, including per capita library collections, should be steadily improved, the equalization and universal accessibility of public cultural services should be promoted, and the conversion of public cultural resources into residents' actual cultural consumption should be improved.

7.5 Promotion of Coordinated Regional Development and Narrowing the Gap in Cultural Consumption

In view of the unbalanced development of cultural consumption and online retail among the eastern, central and western regions, interregional cooperation in the digital economy and cultural industries should be strengthened, digital platforms and logistics resources should be guided to favour underdeveloped areas, the regional disparity in cultural consumption levels should be gradually narrowed, and the coordinated development of cultural consumption should be realized.

References

- [1] Yuan, H. B., & Guo, F. L. (2013). Research on influencing factors of cultural consumption. *Science & Technology Vision*, (5), 33-34.
- [2] Yang, J. L., Li, X. X., Hu, Y. H., & Wu, L. P. (2024). Can digital literacy promotion promote rural residents' cultural consumption growth? *Journal of China Agricultural University*, 29(2), 261-272.
- [3] Wang, K. (2024, March 2). China's online consumption continues to improve. *People's Daily Overseas Edition*, 001.
- [4] Wang, G. S., & Wang, J. (2021). Economic analysis of online retail. *Journal of Beijing Technology and Business University (Social Sciences Edition)*, 36(1), 105-113.

- [5] Gu, J., Shu, H. J., & Tian, X. Y. (2024). Research on the influence mechanism of digital intelligent platforms in narrowing the urban–rural cultural consumption gap. *Rural Economy*, (5), 56-64.
- [6] Han, J., & Liu, Z. Y. (2024). Empowering and boosting residents' consumption through digital-real integration: Mechanism and promotion strategy. *Economic Review Journal*, (3), 60-69.
- [7] Qi, M. Y., & Ge, H. P. (2024). The impact of digital finance on rural consumption upgrading against the background of rural revitalization. *Productivity Research*, (3), 109-114.
- [8] Song, C., Zhong, J., & Li, M. (2024). The impact of coordinated development of logistics and rural e-commerce on rural residents' consumption upgrading. *Journal of Commercial Economics*, (5), 104-108.
- [9] Zhao, Y. M., & Yu, Q. Y. (2024). The impact of platform economy development on rural residents' consumption potential. *Journal of Commercial Economics*, (4), 42-46.
- [10] Zou, S. Z. (2024). The impact of digital economy development on residents' consumption preferences. *Modern Business*, (2), 46-48.
- [11] Yu, Z. P., & Chen, L. (2024). How do new digital infrastructure affect residents' green consumption? *Journal of Technical Economics & Management*, (1), 59-66.
- [12] Xie, H., & Liu, X. W. (2024). Research on the impact of digital inclusive finance on residents' consumption level. *Business Observation*, 12(1), 57-61.
- [13] Xiao, Z. L., Yang, Q., & Miao, Y. Z. (2024). Research on the role of digital technology in promoting residents' consumption upgrading. *Economist*, (1), 17-18.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Research on the Influencing Factors and Optimization Strategies of Enterprise Capital Structure

Chenyu Wang*

School of Economics and Management, Dalian Jiaotong University, Dalian 116041, China

**Corresponding author: Chenyu Wang.*

Abstract

At present's enterprise's financial Management System, The Capital Structure at the heart of its entire process and plays an essential role in Business Efficiency and Profitability. A reasonable capital Structure will be able to optimise resource allocation in an enterprise, lower financing expenses for enterprises, make full use of financial Leverage in practice to increase the Profit level of the enterprise. Given that changes in the economic Environment and an improved Capital Market have made it necessary to make scientific Adjustments to capital Structure as a means to ensure enterprise profitability. This paper combines the actual development of enterprises, discusses the problems in capital structure, analyzes common influencing factors, and makes scientific optimization strategies, so as to support the long-term healthy development of enterprises on the basis of controlling financial risk.

Keywords

enterprise capital structure, influencing factors, optimization strategies

1. Introduction

As the continuous development of the market economy has altered the Financing Environment and Competitive Pattern for Enterprises in their Business Process. In light of the new development Environment for enterprises, it is necessary to improve product Quality and Market Competitiveness while strengthening Operational Efficiency with scientific Financial Management means in place. The primary responsibility for financial decisions at companies; Capital Structure directly affects the efficiency of using corporate funds. In the actual operation of enterprises, there are still many problems in capital structure management. Many companies maintain an excessively conservative capital structure or fail to fully utilise the financial-leverage effect; as a result, their own-profit increase is not substantial enough. Scientifically adjusting capital structure and giving play to its positive role in improving profitability can give great help to optimizing the financing structure of enterprises and ensuring the efficiency of fund use [1, 2].

2. Problems in Enterprise Capital Structure

2.1 Imbalance Between Debt and Equity Financing Ratios

The unreasonable ratio of debt to equity capital in the enterprise's capital Structure has become one of the most serious problems; It will also affect financial Stability and profitability directly. Some enterprises overuse the tax shield effect brought by debt financing and ignore the matching relationship between debt scale and debt-paying ability. In a cycle of rising interest rates, financing costs rise quickly and further squeeze the profit space of enterprises. In addition, the capital Structure's flexibility is weak. Because many enterprises are worried about maintaining control and equity dilution after reducing the proportion of equity financing and relying more heavily on internal accumulation and debt financing for development.. At this time, due to a lack of funds, there will be limitations on investment size. There is also no system for dynamic adjustment available at present. For example, in the expansion stage, they fail to introduce equity investment in a reasonable way. The "static management" mode makes the capital structure out of line with the real development needs of the enterprise, makes it hard to achieve optimal allocation, and affects overall financial performance.

2.2 Relatively Single Financing Channels

Having a single financing channel is one of the common problems in the capital structure of some enterprises, and it is also a main factor restricting the ability to get funds and the level of risk diversification. First, some enterprises depend too much on traditional financing channels. When the credit policy of financial institutions becomes tighter and the credit limit is restricted, the financing ability of enterprises drops greatly, and problems of a tight capital chain can easily appear. At the same time, enterprises are in a passive position in interest rate negotiations, and it is hard to control financing costs. Second, the use of direct financing channels is not enough. Because of weak information disclosure ability, low credit ratings, and other factors, small and medium-sized enterprises cannot smoothly enter the capital market, which is not good for improving capital structure. Third, new financing methods are not used enough. In the new era, innovative financing tools such as supply chain finance and asset securitization are developing step by step, but many enterprises do not know these tools well and do not have enough ability to use them, so their advantages are hard to show. In addition, there is little coordination among different financing channels, and it is hard to build a financing system with a reasonable structure.

2.3 Relatively Weak Financial Risk Control Mechanism

In the process of capital structure management, the imperfect financial risk prevention and control mechanism makes enterprises lack response ability when facing market fluctuations, and this increases the uncertainty in operation. First, the ability to identify and monitor debt risk is not enough, and there is no full analysis of debt scale, maturity structure, and debt-paying ability. At the same time, not enough attention is paid to the problem of an overly high share of short-term debt, and breaks in cash turnover can easily happen. Second, the ability to manage interest rate risk and market risk is weak. Enterprises fail to properly use interest rate derivatives to lock financing costs, which causes unstable financial expenses. Their sensitivity to macroeconomic changes and policy adjustments is also not enough, so it is hard for them to adjust financing strategies. Third, the internal control and budget management system is still not perfect, and the building of related constraints and supervision mechanisms is relatively lagging behind, so the efficiency of fund use is low. In addition, the low transparency of financial information can also affect later decisions of managers.

2.4 Weak Coordination Between Capital Structure and Strategic Investment

Capital structure is also closely related to enterprise development strategy, but from the view of specific management work, the problem of weak coordination is quite clear. First, there is a serious mismatch between fund allocation and development direction. After finishing financing work, some enterprises cannot get funds back in a short time, and the funds are spread over many low-efficiency projects and non-core businesses. The efficiency of resource allocation is low, and the lack of clear investment priorities also adds risk factors in development. Second, the building of evaluation mechanisms is relatively lagging behind. It mainly depends on experience judgment and short-term return expectations, which can easily lead to deviation in investment decisions. Third, financial flexibility is not enough. A highly leveraged structure limits the room for refinancing, and in a market downturn it is hard to quickly adjust investment strategies.

Enterprises face big obstacles in strategy implementation, which is not good for the improvement of long-term value.

3. Influencing Factors of Enterprise Capital Structure

3.1 Financing Channels

As an important way for enterprises to obtain funds, financing channels have an important influence on capital structure. First, whether financing channels are easy to get will affect the choice of capital structure. In the real environment, enterprises with higher credit ratings and larger information disclosure are more likely to get low-cost loans, which greatly raises the share of debt financing. Enterprises with limited financing channels usually rely on internal accumulation and equity financing, and this difference causes clear division in capital structure among different enterprises. Second, the difference in financing costs guides the allocation of capital structure. In later decisions, enterprises mostly compare the capital costs of several channels and look for a balance point between optimal cost and controllable risk. The cost feature is an important variable affecting capital structure. Third, the institutional environment of financing channels is closely connected with the level of market development. In an environment with a sound financial system and a developed capital market, enterprises can choose from many financing methods, which is good for building a diversified capital structure. In addition, regulatory policies, information disclosure, and market access conditions also indirectly affect capital structure.

3.2 Enterprise Size

Enterprise size is one of the internal factors affecting capital structure. Because of differences in financing ability, risk-bearing level, and other conditions, enterprises of different sizes show different capital structure features. First, larger enterprises usually have stronger financing ability and can get long-term funds at a lower cost. At the same time, these enterprises can also enter the capital market more easily and use share issuance or bond issuance for diversified financing. Second, it is related to risk-bearing ability. Large enterprises have stable cash flow and stronger ability to resist risk, so they can bear a higher financial leverage. Small and medium-sized enterprises have larger operating fluctuations. In financing decisions, they are usually more careful and tend to control debt size in order to reduce financial risk. Third, management and governance structure also change with enterprise size. Large enterprises have relatively sound internal control systems and professional financial management teams, which makes the dynamic adjustment and scientific planning of financing structure easier. Small enterprises are a bit weaker in this aspect, and the forward-looking nature of capital structure decisions lacks strong support.

3.3 Financial Risk

Financial risk is a key factor that enterprises should focus on in capital structure decisions. First, debt-paying ability directly limits the level of capital structure. If debt-paying ability is weak, the proportion of debt financing should be reduced to avoid the accumulation of financial risk. When profitability is high and cash flow is stable, the debt ratio can be properly increased to use financial leverage to raise returns. Second, cash flow liquidity affects the choice of financing methods. Enterprises with large cash flow fluctuations usually prefer equity financing and keep a lower debt level, so as to further improve financial flexibility, which makes the financial structure show features highly related to operating conditions. Third, external economic risk cannot be ignored. Under economic cycle fluctuations and interest rate changes, the level of financial risk faced by enterprises will also change. In later decisions, enterprises should look at the whole situation, fully consider the macro environment, and scientifically adjust the financing ratio and structure.

3.4 Industry Characteristics

Industry characteristics are an external factor influencing enterprise capital structure. First, the asset structure of an industry affects the choice of financing methods. For capital-intensive industries such as manufacturing and electric power, assets have higher collateral value, so it is easier to get support from debt financing. Light-asset industries such as the internet and services mostly rely on intangible assets and human capital, so they more often choose equity financing. Second, the profit model of an industry also has an

opposite effect on capital structure. In most cases, industries that are in a mature stage and have stable demand have smaller income fluctuations and more stable cash flow, which can support a higher level of debt financing. Third, the level of industry competition and the policy environment also play an important role. In many industries, competition keeps increasing, and enterprises need to continue technical input and market expansion. This requires high flexibility of funds and also affects the choice of financing methods. Some industries are under strong policy regulation, such as the energy field, and their financing channels and capital structure will also be influenced by policy. Overall, under the joint effect of industry environment and policy factors, the main features of enterprise capital structure are shaped [3].

4. Optimization Strategies for Enterprise Capital Structure

4.1 Optimizing the Ratio of Debt and Equity Financing

The first task in optimizing enterprise capital structure is to reasonably allocate the ratio of debt and equity financing and fully bring out the effect of financial leverage. First, improve the decision-making mechanism for the proportion of debt financing. The suitable debt level is quite different across industries and enterprises, so differentiated management plans should be adopted. Capital-intensive enterprises can use the tax shield effect to reduce financing costs. For high-tech and innovative enterprises, the debt ratio should be strictly controlled and equity financing should be the main choice, so as to avoid risks caused by too much financial pressure. Take an artificial intelligence technology enterprise as an example. This enterprise is in a development stage with high investment in technology research and development and a long profit cycle. If it depends too much on debt financing, it will face large cash flow pressure. Therefore, in the design of its capital structure, this enterprise mainly uses equity financing, and brings in venture capital and strategic investors to obtain research and development funds. Its asset-liability ratio is controlled below 30%. The enterprise only introduces low-cost bank loans in a moderate way when cash flow is relatively stable, and uses them to expand computing power infrastructure, so as to achieve a good match between capital structure and innovative development. Second, optimize equity financing. On the premise of keeping space for profit growth, moderately increasing equity capital helps ease the debt burden and ensure the stability of enterprise operation. It can also help the enterprise build a good image in the capital market, improve the shareholder structure, attract more long-term strategic investors, and promote the long-term healthy development of the enterprise. Third, improve the dynamic adjustment mechanism for optimizing the ratio of debt and equity. Evaluate the ratio of debt and equity according to a reasonable cycle, and flexibly adjust it with profit targets, investment plans, and market environment. During the process of maintaining profitability, enterprises should take flexible strategic policies.

4.2 Promoting the Building of Diversified Financing Channels

A single financing channel can very easily cause concentration of capital risk and limit the flexibility of enterprise financing. Developing diversified financing channels can further reduce financing costs and keep a stable supply of funds under different market environments. First, broaden traditional debt financing channels, including bank loans, corporate bonds, and the issuance of short-term notes. At the same time, combine them with innovative financing methods such as supply chain finance and financial leasing. A manufacturing enterprise, during equipment upgrading, used supply chain finance and financial leasing together for financing. With the credit support of core enterprises, it obtained supply chain finance support and got working capital at a lower interest rate. It also introduced advanced production equipment through financial leasing, turning a one-time large expenditure into installment payments and relieving capital pressure. Together with bank loans and short-term notes, this formed a financing mix and made the sources of funds more diversified. Second, enrich the ways of equity investment. When capital market conditions allow, raise funds through listing, additional share issuance, and the introduction of strategic investors. For non-listed enterprises, explore private equity financing, angel investment, or venture capital channels to get enough funds. These methods can also bring in management experience and industry resources and greatly strengthen the overall competitiveness of enterprises. Third, channel construction should work together with internal governance and risk control systems. Different financing channels involve different financial risks in use, such as interest rate fluctuations, equity dilution, and liquidity pressure. Enterprises should build

scientific fund budgets and debt repayment plans, complete the best combination of financing channels, and ensure a long-term improvement in enterprise profitability.

4.3 Improving the Financial Risk Prevention and Control System

In capital structure optimization and the improvement of profitability, financial risk prevention and control is an important task. It helps ensure stable revenue and avoid common risks such as interest rate risk, liquidity risk, and leverage risk. First, improve the debt risk supervision system. In different situations, use methods such as financial analysis, cash flow forecasting, and interest coverage ratio calculation to grasp the debt level and debt-paying ability, so as to avoid excessive debt and prevent crises such as breaks in the capital chain. Make diversified debt response strategies. For example, in an economic downturn, arrange debt refinancing in advance to further reduce the impact of risk. Second, strengthen interest rate and market risk management. Give play to the role of derivative tools such as interest rate swaps, options, and forward contracts, promote risk hedging, and help enterprises lock the cost of funds. In addition, diversify financing regions and introduce multi-currency financing tools, so as to avoid too much concentration of market risk. Third, increase efforts in internal control and budget management, improve financial systems and related control mechanisms, greatly raise the efficiency of fund use and financial transparency, provide strong support in optimizing capital structure, and ensure overall profitability.

4.4 Increasing Support for Strategic Investment

Capital structure adjustment can affect financial indicators, and it can also support strategic investment through optimized resource allocation, further improving enterprise profitability and market competitiveness. First, allocate funds reasonably and pay attention to high-return projects, such as technological innovation, market expansion, and strategic mergers and acquisitions. This can ensure investment returns and also increase the long-term benefits of the enterprise. An electronic information enterprise, in the process of optimizing capital structure, shifted the focus of financing toward technological innovation and market expansion. The enterprise first increased capital input into a core chip research and development project to improve product technology barriers, and then acquired a regional enterprise with channel advantages to quickly open a new market and complete the expansion of market share. At the same time, in investment decisions, it introduced the net present value and internal rate of return models to strictly screen different kinds of projects, making sure that funds went to high-return areas. When the industry was in an upward boom stage, the enterprise moderately increased leverage to expand investment size. During market fluctuations, it kept a relatively high cash reserve to strengthen risk resistance. Second, optimize the investment decision process and build a scientific project evaluation system, including indicators such as cash flow forecasting, investment payback period, net present value, and internal rate of return. The close coordination between funds and investment decisions helps avoid blind expansion and waste of resources, and it is a core measure for improving the overall profit level. When high-quality investment opportunities appear, enterprises should move funds quickly and carry out merger, acquisition, and expansion plans. Third, strengthen enterprise strategic flexibility, improve financial flexibility, and enhance the ability to respond to the market. For example, in a market downturn, enterprises can reduce high-risk investment and increase working capital reserves; in a rising market, they can expand investment size and seize development opportunities. The continuous improvement of enterprise financial flexibility enables capital structure optimization to support the rise of profitability and also ensures the long-term steady development of enterprises [4, 5, 6].

5. Conclusion

Capital structure adjustment has a deep influence on enterprise profitability. A reasonable capital structure is a strong measure for optimizing resource allocation, and it is also important for enhancing the core competitiveness of enterprises. Managers of related enterprises should give full play to their own initiative, focus on the front line of industry development, and start from the actual development situation of the enterprise. They should reasonably choose measures and strategies such as optimizing the ratio of debt and equity financing, promoting the building of diversified financing channels, and improving the financial risk prevention and control system. In this way, they can strengthen the ability of enterprises to respond to market fluctuations, grasp the internal link between capital structure and profitability, keep improving the

financial management system, and promote the double improvement of economic benefits and social value of enterprises.

References

- [1] Zhang, Y., & Bai, Z. (2024). Monetary policy, enterprise operating risk, and capital structure adjustment. *China Journal of Commerce*, (7), 161–164.
- [2] Yuan, Y., Yang, S., Ge, W., et al. (2024). Research on carbon risk and dynamic adjustment of enterprise capital structure. *Financial Theory & Practice*, (4), 96–107.
- [3] Liu, Z., & Li, Q. (2025). Environmental protection tax law, enterprise rent-seeking behavior, and capital structure adjustment. *Friends of Accounting*, (11), 149–156.
- [4] Guo, J., & Zhu, C. (2024). Digital transformation, human capital structure adjustment, and upgrading of the manufacturing enterprise value chain. *Economic Management*, 46(1), 47–67.
- [5] Wang, J., Chen, Y., & Pan, M. (2024). Research on the influence of capital structure adjustment on enterprise M&A behavior. *Journal of Management*, 21(11), 1603–1611.
- [6] Liu, H., Guo, T., & Hu, K. (2025). State-owned capital operation and innovative development under the background of structural adjustment. *Journal of Hunan University (Social Sciences)*, 39(3), 23–33.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Analysis of the Mechanism by Which Investor Sentiment Affects Stock Market Volatility

Junbin Liang*

School of Economics, Jinan University, Guangzhou, China

**Corresponding author: Junbin Liang.*

Abstract

As a core topic in behavioral finance, investor sentiment and its impact mechanism on stock market volatility have always been a research hotspot in the financial field, serving as a key path to explaining market phenomena. Based on relevant empirical research in recent years, this article first systematically reviews the development process of investor sentiment and its impact on stock market volatility. On this basis, it focuses on analyzing the transmission pathways, boundary conditions, and heterogeneous effects of sentiment on stock market volatility. It also collates traditional econometric methods and text analysis techniques driven by large language models, clarifying the differentiated characteristics of investor sentiment under different market environments. Finally, this paper points out the deficiencies and limitations of current research on investor sentiment, and looks forward to the future where artificial intelligence technology empowers the study of investor sentiment towards high-frequency dynamics and multimodal fusion, aiming to provide theoretical references for risk management and investment decisions of market participants.

Keywords

investor sentiment, stock market volatility, market mechanism, behavioral finance, heterogeneity effect

1. Introduction

In recent years, with the rapid development of behavioral finance, it has been widely applied in investment analysis scenarios, breaking the limitations of traditional financial markets centered on the efficient market hypothesis and rational man. It holds significant importance for asset pricing, the development of investment decisions, and their practical applications [1]. Regarding the impact of investor sentiment on stock market volatility, scholars both domestically and internationally have conducted extensive empirical and theoretical research. In terms of sentiment measurement methods, three mainstream systems have emerged: the direct survey method, the indirect indicator method, and the text mining method [2]. At the level of impact mechanism, it reveals the transmission path through which emotions affect stock price volatility through multiple channels, such as investor trading behavior, market liquidity, and information transmission. At the application level, relevant findings have been widely integrated into investment practices and regulatory policy formulation, propelling behavioral finance from theory to application. They point out that investor sentiment plays a pivotal role in stock market volatility, and through stock market overreaction, investor decision-making behavior and information transmission significantly impact stock market volatility [2].

However, existing research often focuses on a single perspective, which has its limitations. On the one hand, there is a lack of systematic analysis on the research context, hot topics, and development trends among different types of investors. On the other hand, with the rapid development of artificial intelligence and large language model (LLM) technology, methods such as sentiment text mining, high-frequency sentiment measurement, and nonlinear modeling are undergoing paradigm innovation [3]. This makes it difficult for existing reviews to cover the latest advancements promptly.

Based on this, this paper systematically reviews relevant research literature, summarizes the current research status from four aspects: background, process, transmission mechanism, and future outlook, and concludes that investor sentiment causes stock market fluctuations by altering market supply and demand relationships and interfering with information. With the rapid development of artificial intelligence and large language models, sentiment measurement is moving towards high-frequency, dynamic, and multi-model development.

2. Background of the Stock Market and Investor Sentiment

Traditional finance is based on the Efficient Market Hypothesis, which assumes that investors are fully rational, are not influenced by external factors when making investment decisions, and that stock market prices accurately reflect the fundamentals of the company. However, the frequent occurrence of asset bubbles in real capital markets, coupled with liquidity depletion and abnormal price fluctuations (such as IPO premiums and small-cap effects) [4], continuously challenges the assumptions of traditional finance. The rise and development of behavioral finance have integrated psychology into economics, pointing out three aspects: bounded rationality, bounded arbitrage, and emotions as systematic pricing factors, which investors often rely on emotional judgments and generate cognitive biases in an environment of uncertainty and information overload. Within this framework, investor sentiment is believed to have a significant impact on stock market volatility, with overconfidence, sentiment clustering effects, optimistic sentiment leading to high-risk investments, and pessimistic sentiment leading to risk-averse behaviors all significantly affecting stock market volatility.

Investor sentiment is not merely a simple accumulation of individual sentiments, but rather a combination of various factors such as trading behavior, information dissemination, and national institutional environment. In the stock market, sentiment not only affects the degree of asset pricing deviation but also directly shapes the characteristics of market volatility. Optimistic sentiment drives up valuations and trading activity, while pessimistic sentiment triggers panic selling and liquidity contraction. Based on this, academic research on investor sentiment has gradually shifted from theoretical exploration to empirical quantification, undergoing a development process from theoretical foundation, methodological maturity, to technological innovation.

3. Evolution and Cutting-edge Development of Investor Sentiment Research Methods

3.1 Theoretical Foundation Period: Noise Trading Model

In the early stage of investor sentiment research, DeLong proposed the DSSW model, indicating that noise traders would push up price risk, inhibit arbitrage, cause prices to deviate from fundamentals, and obtain higher returns by bearing self-induced risks, which could explain various financial anomalies. This research has propelled the accelerated development of behavioral finance and laid a foundation for the study of investor sentiment [5].

3.2 Method Maturity: Traditional Econometric Models and Proxy Indicator Methods

With the continuous development of behavioral finance theory and financial econometric methods, scholars have adopted empirical verification methods to test the correlation between investor sentiment and stock market volatility, forming a mainstream research paradigm of “econometric model + proxy indicators”. In econometric models, numerous scholars have incorporated sentiment variables into GARCH-type models, stochastic volatility models (SV), and vector autoregressive models (VAR) to depict the volatility characteristics of the stock market. By introducing sentiment variables into these models, they aim to examine their impact on volatility [6]. Taking the closing prices of the Shanghai Composite Index from February 27,

2007, to May 14, 2008, as an example, we characterized the volatility features and conducted data fitting analysis on the sample data using five types of GARCH models and three types of SV models. It was found that these models could indeed effectively depict the volatility features of the stock market. The research perspective has shifted from a single emotional indicator to predicting market cycles (bull and bear markets) and examining the impact on different investors [4].

3.3 Period of Technological Innovation

Nowadays, with the rapid development of artificial intelligence and large language model technology, text data has become the core source of sentiment measurement, including social media comments, financial news, stock forum comments, analyst research reports, corporate announcements, and so on. The early bag-of-words model has been replaced by today's large language models (GPT), significantly enhancing the ability to understand context and recognize and monitor complex information, pushing sentiment analysis into a new stage. Emotions have a significant impact on stock market return volatility, exhibiting heterogeneous and asymmetric effects. The sentiment of individual investors needs to be captured from social commentary, retail trading data, web searches, etc., to reflect short-term sentiment in the stock market. Institutional investors, on the other hand, need to measure sentiment from analyst research reports, changes in institutional holdings, etc., to reflect the long-term impact on the stock market. This heterogeneity analysis has been applied to future research on quantitative investment, focusing on corpus-based multivariate models and dedicated high-frequency indicators [7]. It has deeply integrated with macro-prudential management, algorithmic trading, and market microstructure, marking a new stage of refined, dynamic, and computable sentiment research. With the continuous improvement of measurement methods, scholars have been able to identify the pathways of sentiment effects more accurately and thus systematically reveal the core transmission mechanism of investor sentiment on stock market volatility.

4. Driving Mechanism of Investment Sentiment on Stock Market Volatility

4.1 Direct Conduction Mechanism

The direct driving mechanism of investor sentiment on stock market volatility lies in its ability to alter the balance between market supply and demand, thereby influencing stock prices and their volatility levels. The transmission logic is clear and intuitive. Specifically, when investor sentiment is optimistic, market demand increases, driving stock prices up; whereas when sentiment is pessimistic, demand decreases, triggering excessive selling behavior and leading to a decline in stock prices. Regarding the driving factors of investor sentiment, academia has drawn varying conclusions from different research perspectives [12].

4.2 Indirect Conduction Mechanism

Investor sentiment not only directly affects stock market volatility, but also exerts influence on stock prices and the market through three indirect channels. Firstly, liquidity channel - optimistic sentiment enhances market activity and liquidity, reducing volatility; extreme optimism or pessimism leads to liquidity depletion, exacerbating sharp rises and falls [2]. Secondly, information transmission channels - emotions affect the efficiency of information interpretation and dissemination. When optimistic, negative information is ignored, while when pessimistic, negative signals are amplified, leading to overreaction [8]. Thirdly, arbitrage restriction channels - price deviations driven by sentiment will increase arbitrage risks, deterring arbitrageurs from entering the market, preventing prices from quickly returning to fundamentals, and perpetuating volatility [9].

4.3 Heterogeneity Effect

Existing research has generally confirmed the driving effect of investor sentiment on stock market volatility, but this influence is not uniform, but rather varies depending on market conditions, investor groups, and company characteristics. With the deep penetration of digital technology and the widespread adoption of ESG (Environmental, Social, and Governance) principles, the degree of digital transformation and ESG performance of enterprises has become an important dimension for evaluating company quality. Both can mitigate stock market volatility through emotional channels. However, existing research primarily focuses on

static analysis, with insufficient exploration of the heterogeneity of economic cycles, industry differences, and ESG sub-categories (environment, society, governance). Cross-national comparisons and endogeneity handling need to be strengthened.

In terms of the heterogeneity of market environments, the impact of sentiment on stock market volatility varies across different market conditions: In terms of the differences between bull and bear markets, optimism in a bull market drives stock prices up, leading to market overheating; in a bear market, pessimism significantly magnifies the decline in stock prices, making it more prone to market panic. In the differences in market maturity, investor sentiment has a more significant impact on stock market volatility in emerging markets, where individual investors are the mainstay and irrational trading is more prevalent. In mature markets, institutional investors are the mainstay, and sentiment has a relatively weaker impact on stock market volatility. Yuan Shuyan and Peng Xiqiang took listed companies in Shanghai and Shenzhen stock markets from 2011 to 2021 as samples and found that good ESG performance can alleviate financing constraints by reducing information asymmetry and transmitting positive signals, while the surge of investor sentiment will further strengthen this effect [10]. However, the study remains confined to static analysis and fails to explore heterogeneity across different economic cycles and industry categories. In addition, Negi and other scholars explored the relationship between investor sentiment, market volatility, and the returns of the Indian ESG index, finding that sentiment and volatility significantly affect ESG performance, with a bidirectional causal relationship, providing empirical support from the perspective of behavioral finance. However, this study has several limitations, as it does not cover the global market [11].

In terms of the heterogeneity of the investor group, the impact of sentiment on stock market volatility varies. Individual investor sentiment can be captured through channels such as social media comments, retail trading data, and online search behavior, reflecting short-term sentiment changes in the stock market; institutional investor sentiment is measured based on data such as analysts' research reports and changes in institutional holdings, and its impact is more focused on the long-term trend of the stock market. These types of heterogeneous analysis results have been widely applied in quantitative investment practices. Deep learning models represented by LSTM and CNN, as well as hybrid models such as CEEMD-CNN-LSTM, have provided new tools for stock prediction and portfolio construction. Relevant research aims to elucidate the application logic of deep learning models in stock analysis [12]. In addition, company characteristics such as ESG performance and degree of digital transformation also affect the transmission effect of investor sentiment on volatility. Good ESG performance and digital transformation can reduce information asymmetry and weaken the impact of investor sentiment on stock price volatility. However, existing studies often fail to distinguish the independent roles of each sub-item of ESG (environmental, social, and governance) and do not incorporate key variables such as digital transformation into their analysis, making it difficult to comprehensively depict the regulatory mechanism of company characteristics on sentiment transmission.

5. Future Development and Challenges of Investor Sentiment Research

Despite the fruitful achievements in emotion research, there are still various deficiencies in current related research as a whole: there are defects in measurement methods, traditional text mining is prone to contextual bias and lag, and dynamic high-frequency emotion indicators have not been widely applied. The integration of research mechanisms is relatively lacking [7], with most research focusing on single-channel analysis and lacking systematic exploration of multi-channel interactive effects and the endogenous evolution process of emotions. The research on heterogeneity is insufficient, with inadequate coverage of analyses on differences among retail investors, institutional investors, and quantitative investors, as well as differences between bull and bear markets, different industries, and countries. The emerging research perspectives are relatively weak, and there is limited research combining large language models, digital development, and ESG directions [11]. Most models are set in a linear manner and have weak endogenous processing capabilities. Meanwhile, there is a lack of research on the transformation of research into practice, and research on the implementation and application of real-time monitoring, forward-looking predictions, and related policy tools is relatively scarce.

Future research needs to construct a dynamic sentiment measurement system that integrates large language models, multimodal data, and high-frequency trading data to achieve methodological innovation and enhance the accuracy and timeliness of measurements [7]. Theoretically internalize emotions, deeply explore the generation, dissemination, and evolution mechanisms of emotions within the market, and integrate

interdisciplinary perspectives such as neurofinance, cognitive psychology, and neurology [12]; expand heterogeneous research, systematically examine the differences in emotional effects under various contexts such as investor structure, market cycles, industry characteristics, and institutional environments; promote frontier integration, deepen the interaction mechanisms between environmental, social, and governance (ESG), digital transformation, and investor sentiment, and conduct nonlinear and high-dimensional empirical research. At the same time, regulators need to focus on norms and ethics, pay attention to data privacy and algorithm interpretability, establish a transparent and compliant sentiment analysis framework, and foster rational investment awareness through investor financial education to reduce market volatility driven by emotions and lay a solid foundation for long-term market stability [9].

In summary, investor sentiment, as a key factor affecting stock market volatility, still has vast room for further research. In the future, it is necessary to continuously deepen research at the methodological, theoretical, and application levels to better serve market stability and investment practices.

6. Conclusion

In summary, investor sentiment is a key variable in behavioral finance and a crucial factor influencing stock market volatility. This article reviews relevant research over the past decade and finds that investor sentiment primarily drives stock market volatility through three pathways: altering market supply and demand relationships, affecting market liquidity, and disrupting information transmission. This volatility exhibits significant heterogeneity and asymmetry. With the rapid development of artificial intelligence, large language models, and text mining technology, sentiment measurement has become more frequent, dynamic, and multi-model, contributing to future investment decisions and risk warnings. However, current research faces deficiencies in measurement accuracy and heterogeneity analysis, necessitating further refinement from an interdisciplinary perspective. Simultaneously, there is a need to promote the implementation of emotion monitoring tools and regulatory policies to ensure more stable market development.

References

- [1] Liu J C, & Mao L G. Research on the Influence Mechanism of Investor Sentiment on Stock Market Volatility from the Perspective of Behavioral Finance. *China Collective Economy*, 2025, (23): 105-108.
- [2] Baker M, & Wurgler J. Investor Sentiment in the Stock Market. *Journal of Economic Perspectives*, 2007, 21(2): 129-152.
- [3] Li X H. A Review of the Impact of Investor Sentiment on Stock Market from the Perspective of Text Big Data. *Financial Customer*, 2025, (12): 34-36.
- [4] Naidoo T, Moores-Pitt P, & Muzindutsi P F. Analysing investor sentiment and stock market volatility of the JSE size-based indices: a GARCH-MIDAS approach. *Risk Management*, 2025, 27: 14.
- [5] Qian C, Qinyun L, & Liang C. Risk asset pricing model and its application in the Chinese stock market based on noise traders. *Systems Engineering — Theory & Practice*, 2010, 30(3): 385-395.
- [6] Qi X Z, Ning Z, & Qin M. Economic policy uncertainty, investor sentiment and financial stability - an empirical study based on the time varying parameter-vector autoregression model. *Journal of Economic Interaction and Coordination*, 2022, 17(3).
- [7] Ma Y Y, Liu Y Z, Liu C L, et al. Multi-angle Public Opinion Analysis of Chinese Investors and Its Role in Stock Market Prediction. *Journal of Northeastern University (Natural Science Edition)*, 2022, 43(08): 1201-1208+1216.
- [8] Cornell B, Landsman W R, & Stubben S R. Accounting Information, Investor Sentiment, and Market Pricing. *SSRN Electronic Journal*, 2014, 2(5): 1-8.
- [9] Tan X Y, Zhang Z L, Zhao X J, & Wang C X. Investor sentiment and limits of arbitrage: Evidence from Chinese stock market. *International Review of Economics & Finance*, 2021, 75: 577-595.
- [10] Yuan S H, & Peng X Q. Can Corporate Digital Transformation Improve Investor Sentiment? — Evidence from Shanghai and Shenzhen Listed Companies. *Statistics and Management*, 2026, 41(03): 105-119.

- [11] Negi P, Kushwah S V, Jaiswal A, et al. Investor sentiment, market volatility, and ESG Index dynamics: an empirical analysis. *Cogent Economics & Finance*, 2025, 13(1).
- [12] Rezaei H, Faaljou H, & Mansourfar G. Stock price prediction using deep learning and frequency decomposition. *Expert Systems with Applications*, 2021, 169: 114332.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Research on the Sustainable Development Capacity of CATL - Based on Financial Data from 2015 to 2024

Yukang Liang*

Management College, Hebei University, Baoding 071002, China

*Corresponding author: Yukang Liang.

Abstract

This paper selects the leading company in the battery power industry, CATL Company, as the sample. By using the company's financial data from 2015 to 2024 and that of 10 companies in the same industry for internal comparison, the DuPont analysis method is taken as the core. Combined with the sequential substitution method, the impact degree of the three factors - net profit margin, total asset turnover rate, and equity multiplier - on the return on equity (ROE) is quantitatively calculated. The sensitivity of the indicators is verified through SPSS empirical analysis, and a comparative study is conducted with the leading company in the industry, BYD. Through the triple analysis of industry horizontal comparison, self-vertical comparison, and double-leading company comparison, the financial driving logic of CATL Company is dissected. Its current shortcomings are identified and optimization strategies are proposed. The research concludes that CATL Company has the financial characteristics of “profit-driven and stable leverage ratio”, and determines the financial differentiation characteristics between CATL Company and BYD Company. This also provides certain reference value for financial analysis and optimization of enterprises in the battery power industry.

Keywords

DuPont analysis method, CATL, battery power industry, return on equity (ROE), financial characteristics

1. Introduction

With the rapid development of the global new energy industry, the new energy battery power industry, as a strategic emerging industry, is highly competitive. The financial health and profit-driven efficiency of enterprises have become the key indicators to measure the core competitiveness of the new energy battery power industry. As the leading enterprise in the new energy battery power industry, CATL has a financial operation model, profit quality and risk control ability that are not only crucial for its own sustainable development but also have a benchmarking and exemplary effect on the entire industry. Therefore, this article selects CATL as the research object, which is of certain typicality and representativeness.

Compared with the one-sidedness of single financial indicator analysis, the DuPont analysis method can decompose such a comprehensive indicator like ROE into three factors: profit, operation and leverage. It penetrates the source of the enterprise's profitability layer by layer, understands the advantages and disadvantages of financial strength, and makes up for the deficiency of traditional indicator analysis that cannot reflect the driving mechanism. Combined with the chain substitution method, it can quantify the influence

intensity of factors, making financial analysis scientific, rigorous and targeted. From the perspective of industry characteristics, the battery power industry where CATL is located is a capital and technology-intensive enterprise. The matching of profitability, operation and capital structure is very important. The DuPont analysis has the characteristics of analyzing the battery power industry and can comprehensively reflect the true financial situation of the enterprise in all aspects. Therefore, choosing the DuPont analysis method to study the financial characteristics of CATL can provide good inspiration for similar enterprises in the battery power industry. Hence, this article selects the DuPont analysis method to study the financial characteristics of CATL.

2. Literature Review

2.1 Current Research Status

2.1.1 Theoretical and Applied Research on DuPont Analysis Method

As a classic financial comprehensive analysis tool, the DuPont analysis method starts from the return on equity (ROE), and decomposes it into three major factors: net profit margin, total asset turnover rate, and equity multiplier. It reveals the intrinsic interrelationship among the three aspects of a company's profitability, operational capability, and debt-paying ability. In terms of theoretical improvement, Zhang and Wang (2014) pointed out that the existing DuPont system has significant flaws such as indicator fragmentation, failure to consider profit quality and cost structure. Based on this, they logically and clearly proposed an improved DuPont analysis framework, which has excellent applicability for manufacturing enterprises [1]. Later, Xu (2020) made a very solid optimization of the indicator system, adding the examination of cost control and asset efficiency, truly making the analysis more closely related to the operating reality of real enterprises [2]. From the perspective of industry application, Wen et al. (2023) applied the DuPont analysis method to the financial industry and manufacturing industry, systematically and rigorously testing its effectiveness in financial diagnosis in different fields [3]. Zhang and Hu (2021) reconstructed the DuPont research framework from the perspective of enterprise activities, and concluded that the profit factor and operational factor have a dominant driving effect on the ROE of manufacturing enterprises [4]. Finally, Zhang and Zhou (2019) constructed a suitable DuPont financial ratio model for the manufacturing industry based on an automotive manufacturing enterprise as a sample, which is an excellent reference example for the methodology part of this paper [5].

The application of the DuPont analysis method in the field of new energy and battery power has achieved certain application: Zhang and Zhu (2021) started from the Harvard analysis framework and used DuPont factors to conduct a financial analysis of battery power enterprises, but it was more inclined to strategic interpretation and lacked sufficient quantitative calculation [6]. Xu and Wang (2022) took CATL as a sample, naturally and appropriately incorporated DuPont indicators into the value assessment system of energy storage enterprises, and rigorously demonstrated the actual impact of profit, turnover, and leverage on enterprise value [7]. From this, it can be seen that the DuPont analysis method is applicable to the battery power industry, but there is no mature analysis paradigm that fits its technical-intensive, heavy assets, and high growth characteristics.

2.1.2 Financial Research on the Battery Power Industry

As the new energy industry is a strategic emerging industry of the country, the financial research on the battery power industry has received considerable attention. Currently, the academic community has three clear and logically coherent research directions in this regard.

From the perspective of policy and industry chain impact, the existing literature holds the view that Jin (2022) first demonstrated that the correlation of the new energy vehicle industry chain is high, so policy adjustments and industrial chain synergy jointly affect the financial performance of enterprises [8]. Subsequently, Zhang et al. (2023) made certain supplements to this, namely that the subsidy reduction policy forces enterprises to improve financing efficiency and optimize capital allocation, and that midstream battery enterprises are the most directly and obviously affected by the policy [9]. This is connected with Rao et al. (2022)'s analysis of the impact of the dual credit points policy on the heterogeneity of financial performance of the upstream and downstream of the industry chain, where the midstream battery link benefits the most [10].

Regarding the relationship between R&D investment and sustainable development, Wu et al. (2022) conducted a very clear and rigorous investigation based on panel data, thereby obtaining the U-shaped

relationship between R&D investment of the new energy industry chain enterprises and enterprise value, and the gradual regulatory effect of government subsidies [11]. Song and Zhang (2023) made excellent supplements from the perspective of technological innovation [12].

Chen (2021) made a very clear and logically coherent analysis of optimizing profit quality from the perspective of light asset model regarding enterprise value and operation management, thus providing excellent reference for battery power enterprises to reduce costs and increase efficiency [13]. A very good complement to this is that Tang et al. (2024) demonstrated that digital transformation is conducive to improving the financial performance of manufacturing enterprises through methods of efficiency improvement and value reconstruction [14].

2.1.3 A Comparative Study of Financial Performance of Leading Enterprises in the Industry

Leading enterprises serve as the benchmarks for industry development and have significant exemplary significance. However, at present, the existing literature still has some deficiencies in the research on leading enterprises: Ye and Zhuang (2022) demonstrated how leading enterprises in the industrial chain can drive the growth of small and medium-sized enterprises through cost reduction and efficiency improvement, and also pointed out that industrial agglomeration and a favorable business environment will strengthen the driving effect [15]. Zhang and Yu (2022) used CATL as a case to systematically and rigorously analyze its growth logic from embedding to building an innovative ecosystem, but they have not yet examined it from a financial perspective [16]. Zhang and Wang (2024) provided a theoretical explanation from the perspective of strategic rhythm on the mechanism of high-growth enterprises, providing support for the financial strategy of leading enterprises [17]. Currently, there are two quite excellent studies comparing the dual leading enterprises: Wang and Zhang (2024) used BYD as a case, clarifying the path for improving turnover efficiency and stabilizing profits through the entire industrial chain model [18]. While Zhang et al. (2022) conducted a text analysis of BYD's financial characteristics, they did not make a direct and solid comparison with CATL [19]. Therefore, existing studies recognize the dual leading positions of CATL and BYD, but have not yet conducted long-term and standardized quantitative decomposition of the differences in financial drivers using the DuPont analysis method and the chain substitution method.

2.1.4 Literature Review

From the existing literature, it can be seen that there have been fruitful results in the fields of DuPont analysis application, financial performance of new energy enterprises, industrial chain and policy impact. However, there are indeed three obvious deficiencies in these areas: First, the DuPont analysis in the battery power industry is mostly qualitative research and lacks long-term, large-sample empirical verification. Second, the current research on leading enterprises mainly focuses on a single case, thus lacking the DuPont decomposition and driving factor comparison of two leading enterprises over several consecutive years. Third, the existing research has not fully combined industry patterns and case characteristics, so the analysis of the advantages, weaknesses and optimization directions of the financial model of CATL is not complete.

2.2 Research Innovation

1. Innovation in research methods: Break through the limitations of the traditional DuPont analysis method's qualitative description, combine it with the sequential substitution method of DuPont analysis, and more accurately calculate the impact of net profit margin, total asset turnover, and equity multiplier on ROE. At the same time, use SPSS to conduct empirical analysis of index sensitivity to test and enhance the accuracy of the research results, and make up for the deficiency of the quantitative lack in the traditional DuPont decomposition system.

2. Innovation in comparison dimensions. Conducted a systematic and hierarchical horizontal benchmarking of two industry leaders, CATL and BYD, using DuPont factor decomposition and sequential substitution measurement methods to clarify the differences between the two companies' "profit-driven" and "efficiency-driven" financial models. Thus, naturally and appropriately filled the gaps in the existing literature regarding the financial benchmarking analysis of leading enterprises.

3. Innovation in industry adaptability: In response to the characteristics of the "capital-intensive, technology-intensive, and long supply chain" industry of battery power, adjusted and optimized the interpretation logic of DuPont analysis indicators, combined them with current industry policy declines, raw

material price situations, etc., and focused on the industry adaptability of financial analysis to avoid the lack of industry adaptability in existing research.

3. Research Design

3.1 Research Approach and Sample Selection

1. This study adopts the DuPont analysis method as the core research framework. The return on equity (ROE) is decomposed into profit factors such as the net profit margin from sales, total asset turnover rate, and equity multiplier to study the profit factors. Using SPSS, empirical checks are conducted on correlations, regression analysis of industry panel data, and sensitivity of financial indicators to provide empirical data for the DuPont analysis, thereby compensating for the subjective nature of a single DuPont analysis.

2. Regarding sample selection, to ensure the comprehensiveness of the research, 11 representative listed companies in the battery power and energy storage industries were selected, including CATL, BYD, Guoxuan Gaoke, Xinhuaanda, Funeng Technology, Yihua Tong, Penghui Energy, Wei Lan Lithium, Nandu Power, Shengyang Shares, and Paien Technology. The research period is from 2015 to 2024.

3. All data are from the annual financial reports of enterprises on the East Finance website. After removing missing values, a total of 108 valid observations were obtained, resulting in industry sample panel data. At the same time, for the time series data of CATL and BYD from 2023 to 2024, the DuPont chain substitution quantitative analysis was used.

3.2 Core Variables and DuPont Indicators Definition

1. Core Dependent Variable (ROE): Return on Equity (%), a core indicator in DuPont analysis, reflecting the comprehensive profitability of the enterprise

2. DuPont Core Decomposition Indicators:

NP (Net Sales Profit Margin, %): DuPont Profit Factor, measures the efficiency of main business profit, reflecting the enterprise's cost control and product premium ability

TAT (Total Asset Turnover Rate, times): DuPont Operating Factor, characterizes the efficiency of asset operation, reflecting the enterprise's asset utilization and turnover level

EM (Equity Multiplier): DuPont Leverage Factor, reflects the level of financial leverage, measures the enterprise's capital structure and risk control ability

3. Auxiliary Analysis Indicators: GP (Gross Profit Margin, %), FER (Operating Expense Ratio, %), used to analyze the quality of profit and details of cost control

3.3 Construction of Regression Model

SPSS was used to conduct a multivariate linear regression analysis to verify the sensitivity and significance of the DuPont three factors in driving ROE, in order to quantify the degree of influence of each factor. The model is as follows:

$$ROE_{it} = \beta_0 + \beta_1 NP_{it} + \beta_2 TAT_{it} + \beta_3 EM_{it} + \beta_4 GP_{it} + \beta_5 FER_{it} + \varepsilon_{it} \quad (1)$$

In this formula, i represents an individual enterprise, t represents the year, β_0 is the intercept term, $\beta_1 - \beta_5$ are the regression coefficients, and ε_{it} is the random disturbance term. Separate regressions are conducted on the industry samples and CATL, and the driving differences of the DuPont three factors are compared to provide empirical evidence for the DuPont decomposition analysis.

3.4 DuPont Core Formula and Chain Substitution Method

This article adopts the traditional DuPont analysis core formula and introduces the sequential substitution method for quantitative calculation, accurately quantifying the influence degree of each factor on ROE.

1. DuPont Core Formula

$$ROE = \text{Net sales profit margin}(NP) \times \text{Total Asset Turnover Rate}(TAT) \times \text{Equity multiplier}(EM)$$

2. The method of sequential substitution calculation

The chain substitution method is used to calculate the impact of a single DuPont factor change on ROE. It is carried out in the order of “profit factor → operating factor → leverage factor”, and the steps are as follows:

- 1 Set the base period indicators: $ROE_0 = NP_0 \times TAT_0 \times EM_0$
- 2 Set the indicators for the reporting period: $ROE_1 = NP_1 \times TAT_1 \times EM_1$
- 3 Alternative net sales margin (NP), maintaining TAT and EM unchanged: $ROE_{NP} = NP_1 \times TAT_0 \times EM_0$, The impact of NP = $ROE_{NP} - ROE_0$
- 4 Alternative to Total Asset Turnover Ratio (TAT), keeping NP and EM unchanged: $ROE_{TAT} = NP_1 \times TAT_1 \times EM_0$, TAT impact on the amount = $ROE_{TAT} - ROE_{NP}$
- 5 Alternative Equity Multiplier (EM), maintaining NP and TAT unchanged: $ROE_{EM} = NP_1 \times TAT_1 \times EM_1$, EM impact on the balance sheet = $ROE_{EM} - ROE_{TAT}$
- 6 Overall Verification: Total ROE change = NP impact amount + TAT impact amount + EM impact amount

4. Empirical Analysis

4.1 Industry Descriptive Statistics

Table 1: Descriptive Statistics for the Industry Sample

Variable	Observation count	Mean value	Standard deviation	Minimum value	Maximum value
ROE	108	11.2322	19.00475	-32.76	130.37
NP	108	3.5903	20.19498	-147.34	93.63
GP	108	21.5796	10.17190	-3.35	50.32
TAT	106	0.66937	0.274807	0.075	1.313
EM	108	2.38406	0.922845	1.176	5.788
FER	99	18.2225	18.79880	3.05	159.80

From the descriptive statistical results of the overall sample of the listed industries, it is clearly observable that the Chinese battery power industry has the distinct characteristics of significant fluctuations in indicators and highly differentiated financial performances from enterprises from 2015 to 2024: the average return on equity (ROE) is 11.23%, but its minimum value is as low as -32.76%, and the maximum value is as high as 130.37%. Therefore, the profit levels within the industry are extremely polarized, with a huge profit gap between leading enterprises and those at the tail. It is worth noting that the average sales net profit margin (NP) is only 3.59%, and the difference between the extreme values is nearly 240 percentage points, indicating that the profitability of enterprises within the industry is extremely uneven.

From the available data, the average gross profit margin (GP) is 21.58%, which is at a medium level, but there are indeed significant differences among enterprises. The fundamental reason lies in the differences in product pricing, cost control, and bargaining power in the industry chain among various enterprises. The average total asset turnover rate (TAT) is 0.67 times, which is generally low, which is highly consistent with the basic characteristics of the battery power industry, such as heavy assets, rapid expansion, and long turnover cycle. The average equity multiplier (EM) is 2.38, currently the leverage level is moderate, but some enterprises exceed 5.0. The capital structure of enterprises within the industry varies greatly, and some enterprises expand through high debt financing. The average period expense ratio (FER) is 18.22%, and the maximum value is nearly 160%, which can be used to determine that the ability to control expenses is the most direct and crucial factor determining the quality of a company's profits, and uncontrolled expenses lead to profit loss.

4.2 Correlation Analysis

To verify the intrinsic relationships among the core indicators of DuPont and to determine the direction and strength of the correlation between each factor and ROE, a Pearson correlation analysis was conducted. The results are as follows:

Table 2: Pearson Correlation Matrix of Industry Variables

Variable	ROE	NP	GP	TAT	EM	FER
ROE	1.00	0.564**	0.450**	0.328**	0.248**	-0.181
NP	0.564**	1.00	0.313**	0.299**	0.048	-0.762**
GP	0.450**	0.313**	1.00	-0.214*	-0.300**	0.220*
TAT	0.328**	0.299**	-0.214*	1.00	0.357**	-0.459**
EM	0.248**	0.048	-0.300**	0.357**	1.00	-0.223*
FER	-0.181	-0.762**	0.220*	-0.459**	-0.223*	1.00

Note: * $p < 0.1$, ** $p < 0.05$.

1. The profitability factor is the core driving force of the industry's ROE. From the overall correlation results of the industry, it can be seen that the profitability factor has the most prominent and stable driving effect on the net asset return rate. The correlation coefficient between the net profit margin (NP) and ROE is 0.564, which is significantly positive at the 1% level. Therefore, profitability is the first core element determining the revenue level of battery power enterprises. This conclusion is highly consistent with the characteristics of the battery power industry, which is technology-intensive, has huge R&D investment, and has obvious differences in product value-added: enterprises with core technologies, patent barriers, and scale advantages obviously have greater pricing power and cost control ability, thus being able to maintain higher profit levels in competition and keeping ROE at a high level for a long time.

At the same time, since the correlation coefficient between gross profit margin (GP) and ROE is 0.450, and it is highly significant at the 1% level, and there is a 0.313 obvious positive correlation with the net profit margin, it can be concluded that the product profit margin is the basis for the final net profit realization. To obtain an acceptable net profit, enterprises must first ensure sufficient profit margin space to cover costs, expenses, and various losses, thus profitability quality is the most fundamental support for the financial performance of battery power enterprises.

2. Operating efficiency has a positive pull on ROE, but the effect is relatively limited.

There is a significant positive correlation of 0.328 between total asset turnover rate (TAT) and ROE, but its impact is weaker than the profitability factor. Therefore, it can be inferred that the overall industry operating efficiency has a very limited actual pull on profitability. The specific reasons can be systematically and clearly summarized into three aspects: First, the battery power industry is a typical heavy-asset industry, with huge fixed asset investment and extremely high specificity, so the asset turnover speed is low and the improvement space is limited. Second, the industry demand fluctuates violently and the order cycle is extremely unstable. Enterprises must maintain a high level of raw material and finished product inventory to ensure stable supply, which occupies funds and reduces turnover efficiency. Third, from the correlation structure, it can be clearly seen that although the total asset turnover rate is positively correlated with the net profit margin, the synergy is not strong. Many enterprises in the industry are in an imbalance state of "fast turnover but weak profitability", failing to simultaneously improve scale expansion and profitability quality, so the actual positive contribution of operating efficiency to ROE is very limited.

3. The leverage factor has a positive effect on ROE, but the overall driving effect is weak.

Since the correlation coefficient between equity multiplier (EM) and ROE is 0.248, which is statistically significantly positive at the 1% level, it can be explained that moderate leverage expansion is conducive to improving ROE, but its effect is still weaker than the profitability and operating factors. This leads to the conclusion that the leverage utilization efficiency of the battery power industry is not high.

The main reasons are reflected in two aspects: The use of debt funds by some enterprises in the industry: First, their debt funds are not invested in efficient production capacity and R&D projects, so the leverage cannot form effective returns. Second, factors such as macro interest rate fluctuations and financing cost differences lead some enterprises to have debt costs higher than the asset return rate, thus generating a "leverage negative effect", which neither improves ROE nor reduces financial burden. The significant positive

correlation between equity multiplier and total asset turnover rate, that is, moderate leverage expansion is conducive to asset expansion, but excessive leverage expansion will cause the asset growth rate to decelerate and turnover to be blocked, and the actual benefit of leverage will also be weakened.

4. Cost control ability has a decisive impact on profitability quality.

From the current analysis, the period expense ratio (FER) is negatively correlated with ROE, and it has a -0.762 highly negative correlation with the net profit margin at the 1% level. Therefore, it can be clearly concluded that the cost control level directly determines the efficiency of profit conversion, and is also one of the most important variables affecting financial performance. Period expenses erode profits throughout the entire process from gross profit to net profit: Uncontrolled research and development expenses will temporarily lower profits, uncontrolled sales expenses will inevitably reduce the profit margin of products, and redundant management expenses will inevitably lower the overall operational efficiency.

The period expense ratio has a very significant negative correlation with the total asset turnover rate. Therefore, cost control and operational efficiency complement each other and promote each other. That is, the more refined, digitized, and intensive the cost control in operation management is, the more fully the resources are utilized, the stronger the cost control ability will be, and the more conducive it will be to profitability. Implementing detailed, digital, and intensive cost control in each of the research, sales, and management stages is the direct way for battery power enterprises to improve profit quality and improve financial performance.

4.3 Multivariate Linear Regression Analysis

4.3.1 Regression Results for the Entire Industry Sample

Table 3: Regression Results (Industry Full Sample)

Variable	Coefficient B	t-value	P-value	Significance	Standardized coefficient Beta
Constant items	-38.558	-6.226	0.000	***	—
Net sales profit ratio (NP)	0.815	6.265	0.000	***	0.894
Gross profit margin (GP)	0.316	1.738	0.086	*	0.168
Total asset turnover rate (TAT)	17.389	3.223	0.002	**	0.250
Equity multiplier (EM)	6.858	4.702	0.000	***	0.333
Operating expense ratio (FER)	0.651	4.628	0.000	***	0.655

Model fit: $R^2 = 0.614$, Adjusted $R^2 = 0.593$, $F = 29.237$, $P = 0.00$.

4.3.2 Regression Results of CATL Sample

Table 4: Regression Results (CATL Sample)

Variable	Coefficient B	t-value	P-value	Significance	Standardized coefficient Beta
Constant items	-58.838	-2.948	0.042	**	—
Net sales profit ratio (NP)	1.902	2.325	0.081	*	0.186
Gross profit margin (GP)	-0.253	-0.369	0.731	—	-0.054
Total asset turnover rate (TAT)	49.448	2.596	0.060	*	0.238
Equity multiplier (EM)	4.410	1.137	0.319	—	0.138
Operating expense ratio (FER)	2.521	3.756	0.020	**	0.644

Model fit: $R^2 = 0.991$, Adjusted $R^2 = 0.981$, $F = 93.021$, $P = 0.00$.

4.3.3 Result Analysis

1. Model Fitting and Overall Synergy

The industry regression model, with an adjusted R^2 of 0.593, can explain nearly 60% of the variation in ROE. Additionally, the overall F value of the model is 29.237, which is highly significant at the 1% level. Therefore, it can be demonstrated that the explanatory variables centered around the DuPont three factors have a very strong explanatory power for the industry's ROE. This naturally verifies the applicability and reliability of the DuPont analysis system in financial research within the battery power industry. In contrast, the adjusted R^2 of the individual regression model of CATL is as high as 0.981, far exceeding the industry level, and it also passes the significance test at the 1% level, indicating that the internal factors of CATL, such as profitability, operation, leverage, and cost, have a very strong synergy. The financial driving mechanism is more clear and stable, and the maturity and standardization of the financial system are superior to the industry average.

2. Profit Factor: CATL has both excellent strength and quality in profit driving

From an industry perspective, the regression coefficient of the net profit margin (NP) on ROE is 0.815, which is highly significant at the 1% level, and the standardized coefficient is 0.894. It is the most core and stable driving factor affecting the industry's ROE, meaning that the profit level is the first essential factor determining the comprehensive income ability of battery power enterprises. The coefficient of the gross profit margin (GP) is 0.316, which is significant at the 10% level, and it has a clear positive supporting effect on ROE. Together, they constitute the two main pillars of industry profit.

From the perspective of CATL, the coefficient of the net profit margin on ROE is 1.902, which is highly significant at the 10% level. Therefore, its profit has a much greater actual driving effect on earnings than the industry average, and the driving strength is outstanding. However, the coefficient of the gross profit margin did not pass the significance test. The reason is quite clear: CATL has a very high gross profit base and excellent cost control, so it can directly and reliably convert high gross profit into net profit. The profit conversion efficiency is extremely high, thus forming a dual advantage of "high gross profit base + strong net profit conversion". Correspondingly, the profit purity and sustainability are at the top of the industry.

3. Operating Factor: CATL has a more prominent effect of asset turnover on ROE

From an industry perspective, it is clearly seen that the coefficient of total asset turnover (TAT) on ROE is 17.389, which is highly significant at the 5% level. Improving operational efficiency is conducive to improving enterprise earnings, but it should be noted that it is limited by factors such as heavy assets in the industry and long expansion cycles.

Due to the coefficient of CATL' total asset turnover on ROE being 49.448, which is highly significant at the 10% level, and its effect is almost three times that of the industry average, the main reason is that CATL has a global production capacity layout, achieves large-scale production, and has an extremely lean supply chain management. Therefore, CATL' production utilization rate is always at a high level, with an extremely low asset idle rate, and the capital turnover speed is very fast. Thus, it can amplify the profit advantage through operational efficiency, and operational efficiency has become the most powerful factor supporting CATL' ROE.

4. Leverage Factor: CATL achieves a stable positive effect of leverage, while the overall leverage efficiency of the industry is low

From an industry perspective, the coefficient of the equity multiplier (EM) is 6.858, which is highly significant at the 1% level. However, the leverage levels of various enterprises in the industry are highly inconsistent. Most enterprises have inefficient debt capital allocation efficiency. Therefore, leverage cannot be reliably and fully converted into profit gains, which also leads to some enterprises experiencing negative leverage effects.

CATL' equity multiplier did not pass the significance test. CATL does not drive growth with high leverage but actively maintains a stable capital structure, with long-term debt as the main form of financing, lower financing costs, and all leverage funds are invested in high-return projects. Therefore, the asset return rate is always higher than the debt cost, achieving a safe and efficient "positive leverage effect". Unlike the common practice in the industry of adopting a high-debt expansion model, CATL uses its leverage more rationally and prudently, resulting in lower financial risks.

5. Cost factor: CATL excels in cost control, and its impact on profit erosion is significantly less than that of the industry.

From an industry perspective, the coefficient of the period expense ratio (FER) for ROE is 0.651, which is highly significant at the 1% level. Therefore, the level of cost control directly relates to the quality of an enterprise's profitability. Excessive costs will erode the enterprise's profits.

From CATL's perspective, the coefficient of the period expense ratio for ROE is 2.521, which is highly statistically significant at the 5% level. Moreover, its erosion of profits is significantly lower than that of the industry. The reason is clear: CATL fully utilizes various means such as scale effects, centralized R&D, and digital management to systematically control the period expense ratio at a low level over the long term. Thus, more gross profits can be converted into net profits.

Comprehensive comparative analysis

Compared with the industry as a whole, CATL has a very mature and effective differentiated high-quality financial model of “high profit-driven, efficient operation amplification, stable leverage support, and meticulous cost guarantee”. Therefore, it can break through various predicaments such as profit differentiation, insufficient turnover, inefficient leverage, and excessive costs that are common in the industry. Moreover, in the context of overall industry fluctuations and uneven financial quality, CATL has more stable profits, more coordinated finances, and better risk control. Thus, it is the most typical and successful demonstration of high-quality financial development in the battery power industry.

5. Comparative Analysis

Based on the DuPont three-factor model and combined with the sequential substitution method for quantitative calculation, from three dimensions: horizontal industry benchmarking of CATL, vertical temporal evolution of the company itself, and in-depth comparison between the two leading companies, a multi-level analysis of the driving logic of CATL's ROE is conducted. This enables precise identification of its financial strengths and weaknesses, and realizes the combination and comparison of qualitative analysis and quantitative calculation.

5.1 Comparison of CATL's Indicators with Industry Averages (DuPont)

To demonstrate CATL's financial position in the battery power industry, we compare the average DuPont core indicators of CATL from 2015 to 2024 with those of the industry, and analyze its relative competitive advantages:

Table 5: Comparison between CATL and Industry Average

Indicator Type	CATL	Industry average	Difference range	Advantages/Weaknesses
ROE (%)	25.15	11.23	+13.92	Significant advantages Core advantage
Net Sales Profit Margin (%)	14.24	3.58	+10.66	Minor shortcomings
Total Asset Turnover (times) Equity multiplier	0.61	0.67	-0.06	Steady and controllable
Period expense ratio (%)	2.87	2.38	+0.49	Significant advantages
Indicator Type	8.58	18.19	-9.61	Advantages/Weaknesses

1. The advantage of CATL Technology's ROE is the result of multi-dimensional synergy. Its average ROE is 25.15%, which is 13.92 percentage points higher than the industry average. The comprehensive profitability of this company far exceeds the industry average. This advantage is not the result of a single factor but is the result of the mutual cooperation and complementarity of various factors such as net profit margin from sales, asset turnover rate, leverage efficiency, and cost control. Compared with the leading A-share manufacturing enterprises, its ROE still remains at the leading position, being a model of high-quality development in the industry and a benchmark for driving the improvement of financial performance in the battery power industry.

2. The fundamental source of the profit advantage lies in technical barriers and scale effects. Therefore, the net profit margin of CATL Technology is 14.24%, much higher than the industry average of 3.58%. It has two supporting points: technology and scale. Technologically, CATL Technology has systematically and prospectively optimized battery materials, structures, and systems, thus forming an extremely difficult-to-replicate technical barrier, and the products thereby obtain stable premiums. In terms of scale, the global production capacity layout and high market share bring cost advantages in all aspects of procurement, production, and logistics, and due to the high-quality customer structure and extremely high collection efficiency, the two factors combine to form a truly impregnable profit moat.

3. Weak turnover is an objective and predictable characteristic of the strategic expansion stage, that is, the total asset turnover rate is slightly lower than the industry average. In essence, it is the inevitable result of the company's long-term capacity expansion and global layout. Since the company aims to seize the future market, it actively and systematically increases fixed assets and production capacity investment, which temporarily dilutes turnover efficiency, but the strategy it adopts of “first layout, then harvest” precisely lays a solid foundation for long-term market share increase and profit stability. CATL Technology's “high profit, stable

turnover” model has strong sustainability and is superior to enterprises within the industry that solely rely on high turnover and low profit for survival.

4. Due to the stable and efficient use of leverage, the risk is controllable, and the return is clear, the equity multiplier is higher than the industry but within a reasonable range. It can be seen that CATL Technology has a stable capital structure. Compared with a large number of high-leverage and low-efficiency enterprises within the industry, it forms an excellent contrast: CATL Technology uses leveraged funds in a planned and focused manner for high-return capacity construction and technological research and development, thereby forming a virtuous cycle from leverage to assets and from assets to profit, and diversified and low-cost financing channels have enhanced financial flexibility.

5. Cost control is itself a manifestation of refined operation. The cost expense ratio of CATL Technology is far lower than the industry, which is a direct reflection of the company's extremely strong management ability: the company reduces R&D expenses through centralized R&D, controls sales expenses through brand and channel advantages, improves management efficiency through digital means, and actively and systematically reduces ineffective expenditures. Therefore, the cost control ability has become the most effective aspect for buffering industry price fluctuations, maintaining high profit levels, and breaking industry cycles.

5.2 CATL DuPont Index Longitudinal Comparative Analysis

As this article starts from the enterprise's development stage, it divides the enterprise's 2015-2024 period into three stages: expansion period, optimization period, and leading period. Thus, it systematically observes the dynamic changes of the DuPont three factors.

The expansion period (2015-2018) is a stage where the enterprise actively and systematically expands its scale and market. The total asset turnover rate dropped from 0.72 times to 0.59 times. Due to various factors such as previous R&D investment, capacity construction, and market expansion at this time, the net profit margin dropped from 12.31% to 9.87%. The equity multiplier rose to 3.15, and the ROE remained at a relatively high level of 18% - 20%. Therefore, this stage can be summarized as “emphasizing scale expansion, with a light focus on operational efficiency”.

The optimization period (2019-2021) saw the full release of production capacity, continuous optimization of product structure, and steady expansion of market share. Thus, the operational efficiency improved significantly. The total asset turnover rate rebounded to 0.68 times. Technological upgrading, cost meticulous control, and scale effect combined led to a rebound in the net profit margin to 13.45%. The equity multiplier dropped to 2.92, and the ROE rose to 22% - 24%. The coordination of the DuPont three factors was excellent, with all factors such as profit, turnover, and leverage developing in harmony.

The leading period (2022-2024) is under the circumstances of intensified industry competition, rising raw material prices, and phased overcapacity of production capacity. The enterprise consolidated its high-profit level by leveraging technological barriers and industrial chain advantages. Therefore, the net profit margin increased to 14.92%. It also coincided with the period of macro environment changes and capacity digestion, so the total asset turnover rate dropped to 0.48 times, the equity multiplier remained at 2.87, and the ROE stabilized at above 20%, forming a mature financial model of “profit dominance, stable leverage, and efficiency as the guarantee”. The development resilience has significantly enhanced.

5.3 CATL and BYD's DuPont Analysis Comparison

5.3.1 DuPont Core Indicator Benchmarking

Select the core financial data of the two leading companies from 2020 to 2024, compare the differences in the three DuPont factors, and reveal the differentiated financial models of the two:

Table 6: Comparison between CATL and BYD

DuPont Dimension	Enterprise	2020	2021	2022	2023	2024	Five-year average	Mode positioning
ROE (%)	CATL	11.27	21.52	24.67	24.04	24.13	21.13	Profit-driven type
Net sales profit margin (%)	BYD	7.43	3.73	16.14	24.40	26.05	15.55	Efficiency-driven type

DuPont Dimension	Enterprise	2020	2021	2022	2023	2024	Five-year average	Mode positioning
Total Asset Turnover (times)	CATL	12.13	13.70	10.18	11.66	14.92	12.52	Significantly leading
Equity Multiplier	BYD	3.84	1.84	4.18	5.20	5.35	4.08	Low
Period expense ratio (%)	CATL	0.390	0.562	0.723	0.608	0.481	0.553	Significantly low
DuPont Dimension	BYD	0.790	0.870	1.074	1.027	1.062	0.965	Significantly leading
ROE (%)	CATL	2.263	3.322	3.397	3.262	2.877	3.024	Steady
Net sales profit margin (%)	BYD	3.119	2.837	4.068	4.516	3.943	3.697	High
Total Asset Turnover (times)	CATL	11.26	7.39	9.55	6.99	7.69	8.58	Significantly leading
	BYD	11.87	14.22	10.95	10.12	9.88	11.41	High

1. From the ROE trend, we can observe the differentiation between the two and the different driving logic behind it: The ROE of CATL has remained at a high level and has been steadily increasing. It rose from 20.52% in 2020 to 21.52% in 2021, and then fluctuated slightly and gradually declined in the following years. In sharp contrast, the ROE of BYD has continued to rise. Later on, BYD surpassed CATL. The driving logic of the two is quite clear. CATL maintains a high ROE by generating high profits, while BYD boosts its ROE rapidly through high asset turnover. Thus, it naturally exhibits the distinct financial model differences of “profit-driven” and “efficiency-driven”.

2. The gap in profit levels is essentially caused by the different business structures: The average net profit margin of CATL over the past five years is significantly higher than that of BYD. The fundamental reason is that CATL mainly focuses on battery and energy storage businesses, with high technical barriers and high product value-added, resulting in purer profit quality. In contrast, BYD's business is mainly centered around vehicle manufacturing, with a high proportion of vehicle business and a relatively low net profit rate in the vehicle industry, so its overall profit level has been pulled down. Additionally, CATL has stronger vertical integration capabilities in the supply chain and more prominent cost control advantages.

3. Due to the different industrial chain models, the gap in operational efficiency is very clear: The total asset turnover rate of BYD is significantly ahead of that of CATL. The fundamental reason is that BYD adopts an integrated operation model of vehicle manufacturing and battery, thus having prominent internal synergy effects, and the turnover efficiency of inventory and funds is higher. While CATL mainly focuses on supply business, raw materials and finished product inventories occupy more funds, and the customer payment cycle is also longer.

4. Due to the continuous optimization of leverage levels, each enterprise adopts different prudent strategies, so the equity multipliers of both enterprises are showing a downward trend and are in the stage of actively reducing financial risks. CATL's leverage level is more stable, optimizing the capital structure through its own profits and equity financing, while BYD's leverage scale has declined in recent years, but it is still higher than CATL at present. Therefore, its debt repayment pressure and financial risks are also greater.

5.3.2 Comparison of results from dual-headline sequential substitution method

Using the sequential substitution method, quantitative calculations were conducted for the core indicators of CATL and BYD from 2020 to 2024. The impact amounts and changes of each factor on ROE are summarized as follows:

Table 7: Chain Substitution Analysis for CATL and BYD

Year	Company	Annual ROE Change	Net Profit Margin (NP) Impact	Total Asset Turnover (TAT) Impact	Equity Multiplier (EM) Impact	Equity Multiplier (EM) Impact
2020-2021	CATL	+10.25	+1.84	+3.42	+4.99	Profit, turnover, and leverage drive together, with leverage contributing the most.

Year	Company	Annual ROE Change	Net Profit Margin (NP) Impact	Total Asset Turnover (TAT) Impact	Equity Multiplier (EM) Impact	Equity Multiplier (EM) Impact
	BYD	-3.70	-0.72	+0.21	-3.19	The decline in net profit and the contraction of leverage are the core drag factors.
2021-2022	CATL	+3.15	-3.52	+3.23	+3.44	Turnover and leverage offset the decline in net profit, and the overall performance improves slightly.
	BYD	+12.41	+0.52	+4.31	+7.58	The significant expansion of leverage is the core driver, while turnover and profit assist.
2022-2023	CATL	-0.63	+1.70	-2.32	+0.00	The increase in net profit offsets the decline in turnover, and ROE remains basically stable.
	BYD	+8.26	+0.22	-0.22	+8.26	The significant expansion of leverage is the only core driver.
2023-2024	CATL	+0.09	+6.46	-6.18	-2.76	The increase in net profit offsets the declines in turnover and leverage, and ROE increases slightly.
	BYD	+1.65	+0.70	+0.91	+0.04	Profit and leverage drive together, while turnover slightly hinders.
Five-year average	CATL	+3.22	+1.62	-0.46	+2.06	Leverage dominates, turnover assists, and profit contributes minimally.
	BYD	+4.66	+0.18	+1.30	+3.18	Profit, turnover, and leverage drive together, with leverage contributing the most.

Comparison and analysis conclusion:

CATL is a typical profit-driven benchmark enterprise. Its three characteristics of high profitability, stable leverage, and strong control are very clear. The product premium brought by technical barriers, the cost advantage brought by supply chain integration, and the expense control ability brought by refined management all constitute a complete and powerful core competitiveness. Currently, the large-scale capacity expansion and the cultivation period of the energy storage business occupy a large amount of assets, so the total asset turnover rate is relatively low. This is also the main factor restricting the improvement of ROE at present.

BYD is an efficiency-driven benchmark enterprise. Therefore, its high turnover, strong synergy, and wide layout characteristics are very obvious: The integrated operation of vehicles and batteries brings excellent synergy effects in the industrial chain, and the diversified business layout is conducive to risk dispersion. Therefore, the asset turnover efficiency keeps improving. The excessive proportion of vehicle business has led to the current overall profit level being not high and the leverage level being relatively high. Therefore, financial risk control remains an important issue.

Due to the different strategic positioning and business structure of the two. CATL relies on the battery main business and builds profit barriers through technology and scale, while BYD lays out the entire industrial chain and builds efficiency advantages through integration and collaboration. Therefore, both models are typical representatives of the industry and provide feasible references for enterprises in the same industry. CATL's high-profit model has stronger anti-risk ability, and BYD's high-turnover model has greater growth space.

6. Comparison and Analysis Conclusion

China's battery power enterprises are still in a stage of rapid development and fierce competition. Moreover, the financial conditions of enterprises within the battery power industry vary greatly, with significant differences in performance, overall low operational efficiency, varying levels of benefit control, and low utilization of liabilities. These are common phenomena. At the industry level, profitability and cost control determine the comprehensive financial performance of China's battery power enterprises, and the level of enterprise asset operation and asset-liability structure are key; the direction for enterprises to enhance their core competitiveness.

From the perspective of financial characteristics, CATL has formed a typical financial model characterized by high profit driving, stable leverage support, strong cost control, and relatively weak operational efficiency. CATL demonstrates strong profitability, a relatively stable financial structure, strict risk control, and relatively strong profitability through its leading technology, scale, and industry integration advantages. Its comprehensive strength is more prominent compared to other types of enterprises, but it also has obvious disadvantages such as slow overall asset turnover, operating pressure accompanying large-scale growth, and low asset utilization efficiency. The low operational efficiency has become the main factor restricting the further improvement of its financial performance.

Optimization suggestions for CATL:

1. Capture the current profit characteristics and advantages, strengthen current advantages to continue investment in cutting-edge technology development, maintain a high-profit product structure, rely on its own technological advantages to maintain the advantage of monopolized prices; at the same time, grasp the accumulation of upstream resources, stabilize material costs, and be less affected by price wars and cyclical changes in the industry to maintain the current net profit margin advantage.

2. To comprehensively improve operational efficiency, actively release existing assets, reasonably optimize production capacity layout and asset utilization methods, accelerate the release of new industry capabilities, and strengthen the management of inventory and accounts receivable, shorten the holding period of futures, and accelerate the turnover speed. Use digital and lean operation methods to improve production capacity utilization efficiency and turnover efficiency, and reverse the situation of low turnover rate.

3. Since it is necessary to maintain a stable leverage and capital structure, it is necessary to balance "expansion" and "risk". On the basis of maintaining an appropriate debt level, actively control the proportion of short-term liabilities, lower financing costs, and make full use of abundant operating cash flow to strengthen the company's financial safety cushion. Therefore, an optimal balance point can be found between company expansion and financial balance, forming a high-quality and sustainable capital structure.

4. Based on enhancing supply chain resilience and establishing long-term stable supply chain cooperation and price hedging mechanisms, systematically and plannedly establish a long-term stable supply chain cooperation and price hedging mechanism to reduce the risk of raw material price fluctuations, and actively and precisely control period expenses to truly turn scale advantages into cost advantages, profit advantages.

Reference suggestions for similar industries and other enterprises:

1. Small and medium-sized enterprises: Find a niche market, start by doing well in detail before expanding, first consolidate profitability quality and cost control capabilities, do not blindly expand without considering a decline in turnover and a rising debt ratio; deeply focus on the niche market, take a specialized and professional path, first use technology and efficiency to establish development and survival capabilities, do not engage in head-on competition with leading enterprises in homogenized competition.

2. Mid-sized catch-up enterprises and enterprises in the catch-up period: Learn the development ideas of leading enterprises, fill in the shortcomings in efficiency; draw on the thinking of CATL's industry-acquisition and meticulous management, improve asset turnover and operational efficiency; ensure stable profitability and appropriately utilize leverage to gradually narrow the gap with leading enterprises, scale gap, cost gap, and risk control gap.

3. The entire industry: Take a high-quality and stable development path. Industry enterprises should abandon the low-level competitive mode of low-price competition and embark on a development path

dominated by technological innovation, development, efficiency, and stability; attach importance to product research and development, control the occurrence of risks, and improve cash flow. The industry should shift from rapid development to high-quality development and improve financial health and international competitiveness as a whole.

7. Research Prospects

This paper combines the traditional DuPont analysis with the chain substitution method and empirical testing to improve the financial performance analysis method for battery power enterprises. The conclusions of this article are helpful for enterprise financial diagnosis, strategy formulation, and competitive benchmarking, and are a relatively practical reference. Future work can further expand on this: First, continue to expand the breadth of research on leading overseas enterprises by comparing and analyzing their competitive benchmarking situations; second, add factors such as ESG, R&D investment, and policy impacts to enrich the analysis of financial influencing factors; third, apply more long-term or more frequent data for rolling analysis to observe the changes in enterprise finances; fourth, extend to the research of the entire industrial chain to study the financial transmission and correlation effects between upstream and downstream. Through further research on these issues, more theoretical support and beneficial references can be provided for the healthy development of new energy enterprises, industrial development, and investment decisions.

References

- [1] Zhang, Y., & Wang, X. J. (2014). The Improved DuPont Analysis System and Its Application. *Finance and Accounting Monthly*, 1, 11-14.
- [2] Xu, M. (2020). Brief Analysis on the Improvement and Application of DuPont Analysis System. *Modern Business*, 1, 114-115.
- [3] Wen, S. B., Zhao, X. S., & Li, H. (2023). DuPont Analysis: Interpretation and Application Cases—Application of DuPont Analysis in Financial Industry. *Friends of Accounting*, 12, 150-155.
- [4] Zhang, Y. W., & Hu, M. Y. (2021). Research on DuPont System from the Perspective of Enterprise Activity Division—Evidence from Listed Manufacturing Companies. *Friends of Accounting*, 7, 109-114.
- [5] Zhang, J. J., & Zhou, S. (2019). Construction of DuPont Financial Ratio Model Analysis System—Taking Jinbei Automobile Co., Ltd. (Group) as an Example. *Finance and Accounting Communications*, 17, 109-112.
- [6] Zhang, H. Q., & Zhu, Y. H. (2021). Financial Analysis of Power Battery Enterprises Under Harvard Analysis Framework—Taking EVE Energy Co., Ltd. as an Example. *Commercial Accounting*, 20, 79-81+85.
- [7] Xu, L. F., & Wang, X. J. (2022). Research on Value Evaluation Model of Energy Storage Enterprises—Taking CATL as an Example. *Friends of Accounting*, 15, 18-25.
- [8] Jin, Y. H. (2022). Research on Improving the Level of New Energy Vehicle Industry Chain in the New Development Period. *Economic Review*, 1, 83-90.
- [9] Zhang, H., Yu, Y. Y., & Zhang, X. H. (2023). Research on Financing Efficiency of New Energy Vehicle Industry Chain Under Subsidy Reduction. *Friends of Accounting*, 21, 89-95.
- [10] Rao, Y. B., Xiong, Y. Q., & Xu, W. (2022). Heterogeneous Impact of Dual-Credit Policy on Financial Performance of Upstream and Downstream Enterprises in New Energy Vehicles. *Systems Engineering — Theory & Practice*, 42(9), 2408-2425.
- [11] Wu, J. J., Ren, Y. X., & Zhang, X. L. (2022). R&D Investment, Government Subsidies and Enterprise Value of New Energy Vehicle Industry Chain Enterprises. *The Theory and Practice of Finance and Economics*, 43(6), 140-147.

- [12] Song, J. L., & Zhang, Q. R. (2023). Research on the Impact of R&D Investment of New Energy Enterprises on Sustainable Development Performance—From the Perspective of Technological Innovation. *Economic Problems*, 12, 123-129.
- [13] Chen, H. (2021). Analysis of Light Asset Business Model and Profitability. *Commercial Economics Research*, 11, 110-113.
- [14] Tang, H. T., Xue, Y. W., & Chen, J. (2024). Digital Transformation of Manufacturing Industry and Enterprise Value Enhancement: Core Mechanism, Model Selection and Promotion Path. *Journal of Management Engineering*, 37(2), 81-99.
- [15] Ye, Z. Y., & Zhuang, Z. W. (2022). Leading Enterprises in the Industrial Chain and the Growth of Local Manufacturing Enterprises: Driving Force or Resistance. *China Industrial Economics*, 7, 141-158.
- [16] Zhang, J., & Yu, B. (2022). How Do Start-Ups Build Enterprise Innovation Ecosystems from Embedding? A Longitudinal Case Study of CATL. *R&D Management*, 34(1), 54-70.
- [17] Zhang, X., & Wang, J. J. (2024). Research on the High Growth Mechanism of Enterprises in the Digital Economy Era—From the Perspective of Strategic Rhythm. *China Industrial Economics*, 2, 173-192.
- [18] Wang, X. L., & Zhang, C. (2024). Path and Mechanism of Value Creation in Whole Industry Chain Business Model—A Longitudinal Single Case Study Based on BYD. *Finance and Accounting Communications*, 12, 96-100.
- [19] Zhang, W., Jin, X., & Li, A. (2022). Financial Analysis Based on Natural Language Processing Technology—Taking BYD as an Example. *Friends of Accounting*, 23, 28-36.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

What Makes Consumers Buy: A Study of Go–Stop Psychological Signals in L'Oréal's FMCG Marketing

Jiaxin Shi*

Henan Vocational College of Agriculture, Zhengzhou, China

**Corresponding author: Jiaxin Shi.*

Abstract

The fast-moving consumer goods (FMCG) beauty market, as part of the beauty industry, is marked by short purchase cycles, high brand competition, and consumer behaviour characterized by alternating states of impulsive and deliberate consumption. The need to grasp the underlying psychology of consumer behaviour in the FMCG beauty market has emerged as an essential marketing strategy for global brands. The following review paper aims to investigate the relationship between go-signals, consisting of approach-based psychological drivers such as promotional cues, scarcity appeals, promotion focus, and positive affect, and stop-signals, consisting of avoidance-based drivers such as risk perception, prevention focus, and cognitive resistance, in the FMCG beauty marketing arena. The paper is based on peer-reviewed scholarly articles on consumer psychology, regulatory focus theory, neuromarketing, and brand management and draws on the theoretical frameworks to discuss the documented marketing practice adopted by the global brand, L'Oréal. The gaps in the literature are also discussed, revealing the need for cross-cultural comparative and longitudinal studies on the go-stop signal relationship in the digital FMCG arena.

Keywords

approach–avoidance conflict, regulatory focus theory, FMCG marketing, consumer psychology, L'Oréal

1. Introduction

One of the most psychologically complicated classes of products of the fast-moving consumer goods industry is the global beauty and personal care market, whose value is estimated to be more than USD 580 billion as of 2023 [1]. FMCG beauty products are placed in an interesting psychological place: their unit price is low, symbolically rich, bought as a habit, motivated by aspiration, and advertised to both conscious rational thought and deep emotional desire at the same time. In the case of multinational corporations like L'Oreal the largest cosmetics corporation in the world, it is a question of multi-billion dollars to know the precise mental state that causes a consumer to pick up a product, place it in a cart, or leave it on a shelf.

Finding out what motivates purchase is never the key issue, but rather it is the dynamic tension between forces driving the consumer to make a decision and those drawing them out of it. Such a tension has been theorised in various forms in the psychology and marketing literature: as approach versus avoidance motivation [2], as promotion versus prevention regulatory focus [3], as Systems 1 versus System 2 cognition [4], and as go signals versus stop signals in the larger consumer decision-making process [5]. In a way, each of those

frameworks represents the varying aspect of the same underlying phenomenon that purchasing is not a unidirectional cognitive process but the result of the contesting psychological forces at the conscious and subconscious levels.

In this review paper, three main research questions are covered. To start with, what are the psychological cues that serve as go triggers, i.e. what factors in the environment trigger approach motivation and drive consumers to make a purchase in the FMCG beauty setting? Second, what are the signals that serve as stop triggers i.e. what stimuli trigger avoidance motivation and lead consumers to be hesitate, disengage or reject? Third, what is the relationship between the marketing strategy documented by L'Oreal and the ways it reinforced, capitalized upon or otherwise exploited these psychological mechanisms? To answer these questions, this review combines evidence provided by twenty scholarly and peer-reviewed articles on the regulatory focus theory, neuromarketing, scarcity studies, experiential marketing, and brand psychology, and L'Oreal is used as a prime example to do so.

2. Theoretical Background

The analysis of go-stop psychological signals of FMCG marketing is founded on three key theoretical frameworks. The initial and most basic one is Approach-Avoidance Theory, which was developed within the field of environmental psychology. [2] have determined that environmental stimuli produce emotional states that are based on the dimensions of pleasure, arousal, and dominance, and that the emotional states directly predict approach (engagement, exploration, spending) or avoidance (withdrawal, disengagement) behaviours. This model was applied by Donovan and Rossiter [6] to the retail context, where the authors showed that variables of store atmosphere such as lighting, music, spatial/store layout moderate the emotional state and consequently influence purchasing. This framework was also generalised by Hopkinson and Hogg [5], who indicated that approach-avoidance conflict occurs when a product elicits both positive desire (hedonic appeal) and negative resistance (moral concern, price anxiety or perceived risk).

The second one is the Regulatory Focus Theory (RFT) outlined by Higgins [3, 7] that suggests the introduction of motivational orientation as a major mediator of the approach-avoidance relationship. RFT is the view that people work with either one of two basic self-regulatory systems: a promotion focus, which directs their behaviour towards gains, aspirations and ideal outcomes, and a prevention focus, which directs their behaviour in terms of safety, loss avoidance and normative obligations. More importantly, RFT postulates that marketing messages have the most persuasive effect when they produce regulatory fit that is, when the framing of a message is consistent with the regulatory orientation that the consumer has chronic or situationally induced. In a series of experiments Werth and Foerster [8] showed that consumers rated products positively when the product features were similar to their regulatory focus orientation. In the case of luxury brands like L'Oreal, which are involved in the sale of FMCG beauty products, it means that different messaging architectures could be needed to promote the product to the promotion-oriented and prevention-oriented consumer groups.

The third model is the dual-process model by Kahneman [4], which identifies two systems System 1 (fast, automatic, emotional processing) and System 2 (slow, deliberate, analytical processing). System 1 processes are overpowered in the case of FMCG where purchase cycles are short, and products are low-involvement. This observation has been operationalised by neuromarketing studies demonstrating that the emotional assessments triggered by packaging, advertising and price cues are largely mediated by quick, subconscious emotional judgements that are rapidly triggered prior to conscious thought [9, 10]. These three theories, the Approach-Avoidance Theory, the Regulatory Focus Theory and the two-process cognition, present a comprehensive conceptual jargon to the study of the go-stop architecture of consumer buying.

3. Literature Review

3.1 Go Signals: Approach-Oriented Psychological Triggers

An established body of work analyses the role of emotional and environmental stimuli as go signals that drive consumers into purchase. Taking the PAD (Pleasure-Arousal-Dominance) model proposed by Mehrabian and Russell [2] as a basis, Donovan and Rossiter [6] established that retail settings with a high degree of

pleasure and a moderate degree of arousal stimulated approach behaviours containing unplanned buying as well as time-consuming shopping behaviours.

A complementary explanation of go signals is provided by Regulatory Focus Theory, which is an account of regulatory level go signals. Higgins [3] revealed that promotion-oriented persons are particularly sensitive to the advertising frames which highlight the achievement, aspiration and gain. Specifically, Werth and Foerster [8] demonstrated that consumers in a promotion focus were more likely to prefer product characteristics that are over comfort and consumers in the prevention focus preferred over safety characteristics. In the case of L'Oreal, whose core slogan, 'the reason we are worth it,' is a textbook promotion-based appeal, this study details why the core consumer franchise of the given brand has been historically biased towards the aspirational self-improvement discourse. Peng et al. [11] furthered these studies to online contexts and discovered that promotion-oriented consumers formed larger consideration sets and positively reacted to rich information stimuli, and prevention-oriented consumers favoured focused and concise information.

A special but very powerful type of go signal is scarcity appeals. Barton et al. [12] have systematically reviewed 77 studies across 17 countries and found that both quantity-based (limited stock) and time-based (limited offer) product scarcity signals had a significant positive relationship with purchase intention across product categories and that the effect size was enhanced when consumers were in higher arousal states. Chen et al. [13] also showed that perceived scarcity mediated by urgency to purchase directly in live-streaming e-commerce environments, which are becoming more and more applicable to the digital sales strategy of L'Oreal, and mediated the impact of perceived scarcity by increasing impulse purchase through aroused emotions.

Xia et al. [14] presented direct evidence of the use of go signal in the documented marketing practice of L'Oreal through a case study of the experiential marketing programme of L'Oreal with the Chinese female university students. The authors applied the five SEM experience modules proposed by Schmitt [15] which are Sense, Feel, Think, Act and Relate and discovered that the in-person and online experiential activations of L'Oreal were found to produce strong positive emotional reactions and translated into brand preference and purchase intention.

3.2 Stop Signals: Avoidance-Oriented Psychological Forces

Similar to the go signal literature, there is a rich body of research on the literature that has been conducted in the area of research concerning the forces that produce avoidance motivation and interrupt purchase. Wang et al. [16] examined the processes through which the advertising avoidance is triggered by the perception of risk based on the SOR theoretical framework. They discovered that, the privacy risk, performance risk, time-loss risk, and freedom risk yielded different profiles of negative emotional response, which consequently forecasted advertising avoidance behaviour. This suggests that an identical amount of the objective risk of a product could generate radically different levels of stop signal in the face of an emotional baseline of a consumer- a fact that has major implications on how L'Oreal can position its dermatologically active skincare lines in the marketplace, in which ingredient communication has a paradoxical influence on the development of risk perceptions.

The conceptual model of the approach-avoidance conflict in consumer behaviour by Hopkinson and Hogg [5] indicated that there are four types of product-level stop signals: functional avoidance (product does not work), social avoidance (product conflicts with social identity), moral avoidance (product conflicts with ethical values), and hedonic avoidance (product is associated with guilt or negative self-image). Social and moral avoidance are especially relevant to the case of beauty products. A consumer can want a luxury L'Oreal serum (approach) but she is at cross-purposes knowing that its cost is more than she thinks she should spend on herself (moral-prevention stop signal).

The Regulatory Focus Theory provides a mechanized explanation of prevention-oriented stop signals. Vriend et al. [17] presented a very restrictive empirical division of RFT distinguishing between minimal goals (the prevention-seeking ought self) and prevention strategies (vigilance). They discovered that prevention-oriented vigilance systematically yielded risk averse, conservative choice models and decreased readiness to adopt new items- a crucial factor to the L'Oreal innovation channel, where new formulations must surmount customer habitual inertia. Wang et al. [16] ensured that information overload was negatively applied to prevention-oriented consumers in the digital environment, where processing rich product information was perceived as a threat instead of a reassurance. Combined, these results indicate that stop signals do not simply

occur because there were go signals, but are triggered independently by certain marketing stimuli, especially those perceived as uncertain, confusing or socially or morally dangerous.

3.3 The Go–Stop Interaction and Brand Implications

The most advanced branch of the literature looks at the interactions of go and stop signals as opposed to viewing them as independent additive forces. The somatic marker hypothesis developed by Damasio [18] is used extensively in consumer neuroscience and suggests the attachment of valence tags to later moments of decision-making, based on the emotional memory of a previous experience with a brand, which is a sort of pre-programming of either a positive or a negative response prior to the involvement of deliberative thought. In the case of a well-known brand like L'Oreal, decades of advertising may have created rich somatic marker networks in millions of consumers: the L'Oreal packaging, the hearing of the recognisable slogan, even the touch of a recognisable cream texture can all be automatic go-cues. On the other hand, a negative product experience will create a stop signal which will be just as persistent as the first negative experience [18].

Tversky and Kahneman [19] determined that when consumers are facing uncertainty, they do not make systematic judgments but instead use cognitive heuristics, which are availability, representativeness, and anchoring. This implies that go and stop signals do not process equally in FMCG beauty situations. The discovery that losses are feared about twice as much as the same amount of gain, loss aversion, occurs, predicts that stop cues can be smaller than go cues to have the same effect on behaviour.

The synthesis of the role of scarcity marketing tactics with the overall brand equity, where Hamilton et al. [20] posit that scarcity cues as go signals have no threat of activating reactance and brands with lesser trust capital have a risk of triggering avoidance, is proposed by Roux et al. [21] through the synthesis of the work of Hamilton et al. [20] on the nature of scarcity marketing tactics in the marketing mix. The neural value of the same product was found to be modified by brand identity as in the largest-known fMRI study by Montague et al. [22]: the same beverage caused really different reward circuit activity when the brand label was present or not. In the case of L'Oreal, where the brand base ranges through both entry-point mass-market products and high-end luxury brands like Lancome or Yves Saint Laurent, this observation suggests that the same product formula in different brand packages can evoke oppositely distinct go-stop neural processes in the consumers.

4. Discussion and Synthesis

Considered collectively, the literature reviewed leads to one obvious general conclusion they are not the positive and negative ends of one motivational dimension, but are quite different mechanisms with different neurological bases, different sensitivities to marketing stimuli, and asymmetric influences on behaviour. This has far-reaching consequences on the conceptualisation of the campaign design of such FMCG brands as L'Oreal. The paradigm of the current approach to marketing strategy views persuasion mainly as an issue of the amplification of go signals by appealing imagery, aspirational messages, and campaigns. According to the reviewed literature, this is not enough: an effective campaign should also switch off stop signals, not increase go signals, since loss aversion and negative affect asymmetries imply that an effective stop signal can cancel the collective action of multiple positive stimuli.

Regulatory focus, approach-avoidance and neuromarketing literatures are widely consistent in that emotional signals rather than rational signals prevail in FMCG purchase situations. Kahneman [4], Ariely and Berns [9], and Damasio [18] take similar positions by agreeing that the bulk of FMCG buying decisions are made prior to the involvement of conscious deliberation. This practical implication, that feature considerations by System 2 are less decisive than emotional considerations by System 1, is echoed in L'Oreal long-standing emotional-based advertising campaign which is always based on how the product makes the consumer feel above what it does or doesn't do. The success reported in the case study analysis of L'Oreal because we worth it campaign that has been in circulation since 1973 can be viewed as a masterpiece in chronic promotion-focus activation that the slogan directly appeals to the ideal self of the consumer and that purchasing is a self-affirmation process and not a self-indulgence and the stop signal of prevention mode is neutralised.

The place where the literature review shows some significant tension and contradiction is where scarcity and digital marketing are concerned. Barton et al. [12] and Chen et al. [13] and Zhang et al. [23] all find positive impacts of scarcity on purchase intention, but conceptualisation of Hopkinson and Hogg [5] point to

the risks of over-dependence on scarcity, decreased post-purchase engagement, and diminished quality of brand relationships. The tension is especially relevant to the digital approach of L'Oreal, in which AI-based consumer insights are used to detect trends and determine the most effective time to make an offer. The danger is the risk that hyper-personalised, urgency-driven digital marketing will transform promotion-oriented consumers into prevention-oriented ones in the long run as they become aware and resistant to the manipulative practices a phenomenon that is not yet properly considered in the literature examined.

One of the significant weaknesses of the current literature is that it is necessarily Western and East Asian. The data presented by Xia et al. [14] and Dang et al. [24] is also valuable when it comes to L'Oreal-specific data on the Chinese market, however, there is virtually no research that analyses the functioning of go-stop signal dynamics in the context of African, Latin American, or Middle Eastern consumers. This is not just an academic gap but a commercially essential one considering that in its strategy, L'Oreal has clearly set out one billion new consumers in the emerging markets. Also, most of the studies reviewed are cross-sectional in nature, and such a study is not able to establish the longitudinal relationship between repeated brand exposure, somatic marker formation, and the progressive change in go-stop signal sensitivity that transpires through the lifespan of a consumer with a brand. Both of these limitations should be solved in future researches using multi-cultural longitudinal designs.

5. Conclusion

The marketing behaviour L'Oreal recorded in the form of its iconic promotion-driven slogan and structure of the brand hierarchy, its AI-powered consumer analytics and experiential marketing programmes, demonstrates advanced, though not always explicitly theorised correspondence with the psychological principles defined in the observed literature. The successful ability to master the go signal architecture intuitively and types of architecture can be described as one of the reasons why the brand continues to dominate the market. However, with the new digital conditions, especially live-streaming commerce and hyper-personalised targeted advertising, there are new stop signal hazards, namely, psychological reactance, that have not been properly described in the current literature in the context of FMCG beauty.

The three key contributions of the review include the three finally integrated theoretical frameworks that have hitherto existed in silos; the systematized application of the theoretical framework to the actual marketing strategy of L'Oreal; and the identification of specific gaps that can be filled by future empirical research, such as cross-cultural comparison studies, longitudinal research on signal interaction, and experimental research on reactance threshold in the digital FMCG setting. By taking consumer motivation as a system of competing signals, dynamic, and not a linear role of positive stimuli, researchers and practitioners will be in a better position to create marketing that is not only persuasive but one that is sustainable.

References

- [1] McKinsey & Company. (2023). The beauty market in 2023: A special State of Fashion report. McKinsey & Company. <https://www.mckinsey.com/industries/retail/our-insights/the-beauty-market-in-2023-a-special-state-of-fashion-report>
- [2] Mehrabian, A., & Russell, J. A. (1974). An approach to environmental psychology. MIT Press.
- [3] Higgins, E. T. (1997). Beyond pleasure and pain. *American Psychologist*, 52(12), 1280–1300. <https://doi.org/10.1037/0003-066X.52.12.1280>
- [4] Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.
- [5] Hopkinson, G. C., & Hogg, M. K. (2007). Approach-avoidance conflicts in consumer behavior: Towards a conceptual framework. In R. W. Belk (Ed.), *Handbook of qualitative research methods in marketing* (pp. 325–338). Edward Elgar.
- [6] Donovan, R. J., & Rossiter, J. R. (1982). Store atmosphere: An environmental psychology approach. *Journal of Retailing*, 58(1), 34–57.

- [7] Higgins, E. T. (1998). Promotion and prevention: Regulatory focus as a motivational principle. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1–46). Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60381-0](https://doi.org/10.1016/S0065-2601(08)60381-0)
- [8] Werth, L., & Foerster, J. (2007). How regulatory focus influences consumer behavior. *European Journal of Social Psychology*, 37(1), 33–51. <https://doi.org/10.1002/ejsp.343>
- [9] Ariely, D., & Berns, G. S. (2010). Neuromarketing: The hope and hype of neuroimaging in business. *Nature Reviews Neuroscience*, 11(4), 284–292. <https://doi.org/10.1038/nrn2795>
- [10] Plassmann, H., Ramsøy, T. Z., & Milosavljevic, M. (2012). Branding the brain: A critical review and outlook. *Journal of Consumer Psychology*, 22(1), 18–36. <https://doi.org/10.1016/j.jcps.2011.11.010>
- [11] Peng, M., Xu, Z., & Huang, H. (2022). To each their own: The impact of regulatory focus on consumers' response to online information load. *Frontiers in Neuroscience*, 16, Article 757316. <https://doi.org/10.3389/fnins.2022.757316>
- [12] Barton, B., Zlatevska, N., & Oppewal, H. (2022). Scarcity tactics in marketing: A meta-analysis of product scarcity effects on consumer purchase intentions. *Journal of Retailing*, 98(4), 741–758. <https://doi.org/10.1016/j.jretai.2022.06.003>
- [13] Chen, B., Wang, L., Rasool, H., & Wang, J. (2022). Research on the impact of marketing strategy on consumers' impulsive purchase behavior in livestreaming e-commerce. *Frontiers in Psychology*, 13, Article 905531. <https://doi.org/10.3389/fpsyg.2022.905531>
- [14] Xia, B., Han, M., & Zhu, Y. (2021). Experiential marketing of L'Oréal in China: A case study based on consumer behavior of Chinese female university students. In *Proceedings of the 2021 3rd International Conference on Economic Management and Cultural Industry (ICEMCI 2021)* (pp. 2270–2280). Atlantis Press. <https://doi.org/10.2991/assehr.k.211209.373>
- [15] Schmitt, B. H. (1999). [REFERENCE MISSING — please add full citation]
- [16] Wang, H. J., Yue, X. L., Ansari, A. R., Tang, G. Q., Ding, J. Y., & Jiang, Y. Q. (2022). Research on the influence mechanism of consumers' perceived risk on the advertising avoidance behavior of online targeted advertising. *Frontiers in Psychology*, 13, Article 878629. <https://doi.org/10.3389/fpsyg.2022.878629>
- [17] Vriend, T., Hamstra, M. R., Said, R., Janssen, O., Jordan, J., & Nijstad, B. A. (2023). Regulatory focus theory: Disentangling goals and strategies. *Applied Psychology*, 72(1), 231–267. <https://doi.org/10.1111/apps.12376>
- [18] Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. Penguin.
- [19] Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124–1131. <https://doi.org/10.1126/science.185.4157.1124>
- [20] Hamilton, R., et al. (2019). [REFERENCE MISSING — please add full citation]
- [21] Roux, C., Goldsmith, K., & Cannon, C. (2023). On the role of scarcity in marketing: Identifying research opportunities across the 5Ps. *Journal of the Academy of Marketing Science*, 51(6), 1197–1202. <https://doi.org/10.1007/s11747-023-00956-0>
- [22] McClure, S. M., Li, J., Tomlin, D., Cypert, K. S., Montague, L. M., & Montague, P. R. (2004). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron*, 44(2), 379–387. <https://doi.org/10.1016/j.neuron.2004.09.019>
- [23] Zhang, J., Jiang, N., Turner, J. J., & Pahlevan-Sharif, S. (2022). The impact of scarcity on consumers' impulse buying based on the SOR theory. *Frontiers in Psychology*, 13, Article 792419. <https://doi.org/10.3389/fpsyg.2022.792419>
- [24] Dang, M., Kong, X., & Wang, Y. (2022). L'Oréal Paris consumer loyalty research based on the Chinese market. *BCP Business & Management*, 20, 719–726. <https://doi.org/10.54691/bcpbm.v20i.1056>

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Heterogeneous Analysis of RMI on Urban Consumption Vitality: An Empirical Study Based on Panel Data of 50 Cities

Haoyu Qin*

The School of Statistics and Mathematics, Shandong University of Finance and Economics, Jinan 250014, China

**Corresponding author: Haoyu Qin.*

Abstract

Deepening domestic demand drives high-quality urban development. However, some lower-tier cities blindly introduce high-end brands, causing a spatial mismatch. This creates a severe gap between commercial supply and local purchasing power. Based on the Spatial Mismatch and Bottom of the Pyramid (BOP) theories, panel data from 50 Chinese cities (2022-2024) and spatial data of retail stores are utilized. A “Relative Mass-market Index” (RMI) is innovatively constructed. A Pooled OLS model is adopted to test how supply structure fit affects macro consumption. It is found that this activation effect is highly asymmetric. In lower-tier markets, the “downward compatibility” of mass-market retail generates a strong “filling effect,” significantly increasing per capita retail sales. In mature markets, due to a saturated retail ecosystem, the index adjustment exhibits “elasticity desensitization.” The conclusion confirms that high-quality mass-market retail acts as “new infrastructure” for lower-tier markets. It provides an important decision-making basis for local governments to abandon blindly upgrading commercial districts and to implement city-specific policies to unleash long-tail consumption benefits.

Keywords

spatial mismatch, lower-tier market, relative mass-market index, consumption vitality

1. Introduction

1.1 Background

China is currently focusing on building a new development pattern of 'dual circulation'. Under this macro background, promoting consumption and deepening domestic demand has become the core driving force for the high-quality development of urban development. The traditional commercial markets in first- and second-tier cities are currently gradually reaching saturation. However, China still has vast “lower-tier markets” that remain to be explored. These areas have enormous consumption resilience and consumption vitality, and are at a critical stage of consumption upgrading. In some low-tier cities, high-end brands are blindly introduced, causing spatial mismatch and resource waste. This phenomenon has led to a serious disconnect between the supply of physical businesses and the actual purchasing power of local residents. To face these problems,

which is more effective: simple capital expansion, or the structural fit of commercial supply (the relative ratio of mass-market to high-end formats)?

1.2 Literature Review

Recently, the relationship between retail space evolution and macro consumption vitality has been deeply explored. The relevant research framework mainly follows from “agglomeration size” to “spatial structure match.”

Early research mostly focused on the scale expansion and density of commercial resources. Couture & Handbury proved that the spatial agglomeration of non-tradable services attracts population return and reshapes urban vitality [1]. Baviera-Puig et al. proved that the spatial layout of retail formats is directly driven by local demographics [2]. Wang et al. revealed significant heterogeneity in the spatial clustering of commercial facilities; blind layout easily causes a mismatch between objective supply and local foot traffic [3].

However, with the complexity of urban spatial structure, Scholars are gradually shifting their focus toward the spatial mismatch of resources. The “Spatial Mismatch Hypothesis” was first proposed by Kain to explain unemployment caused by spatial disconnection [4]. Sun et al. revealed the negative impact of housing-job mismatch on urban efficiency in Shanghai [5]. Gobillon et al. systematically summarized the micro-mechanisms of spatial mismatch, pointing out that physical disconnection causes severe market exclusion [6]. This classic analytical framework provides an excellent theoretical anchor for analyzing the detachment between retail networks and local consumption bases.

At the level of supply-demand matching between the commercial supply structure and heterogeneous markets, Schiff confirmed that high-end retail relies heavily on large demand networks in core cities [7]. Tang et al. revealed a significant spatial mismatch between service supply and real local demand [8]. Glaeser et al. pointed out that commercial gentrification accelerates local retail churn, but weakening the loss of the original unique local businesses [9]. London and Hart expanded the BOP theory, noting that transnational models cannot blindly copy the existing models of the high-end market [10]. It is necessary to leverage the advantages of the existing market environment and provide customized solutions that align with the actual purchasing power locally.

Overall, existing literature lays a solid foundation for understanding urban retail agglomeration and spatial mismatch. Most studies focus on total store numbers or job-housing mismatch. Few studies explore the filling mechanism of mass-market retail in lower-tier markets from a supply structure perspective. This paper aims to fill this academic gap through empirical testing.

1.3 Research Content

To fill the above research gap, Spatial Mismatch and BOP theories are used as the core logical thread, and innovatively extends them to the field of physical business supply. Panel data from 50 key cities (2022-2024) are utilized. “Chabaidao” (mass-market) and “Starbucks” (high-end) are selected to construct the RMI. The City Commercial Charm Ranking is used as the standard for city commercial levels. Since offline stores are slow-moving variables, a Pooled OLS with Year Effects is constructed. Cluster-Robust Standard Errors (SE) are adopted to ensure the Reliability of Statistical Inference. A one-period lagged independent variable is introduced in robustness tests to alleviate endogeneity bias caused by reverse causality.

From the perspective of research direction, the research broadens the research boundaries of the 'spatial mismatch effect'. It is found that the activation effect of commercial supply structure on domestic demand is significantly asymmetric. In low-tier cities, mass-market retail presents a strong “filling effect”, significantly boosting the local macro consumption vitality. In mature markets, adjustments to this single structure show “elasticity desensitization.” Standardized mass-market retail is confirmed as “new infrastructure” for lower-tier markets.

2. Theoretical Framework and Hypotheses

Spatial Mismatch and Bottom of the Pyramid (BOP) theories are combined to analyze the internal mechanism of RMI stimulating domestic demand from the perspective of structural adaptation, and propose the corresponding research hypotheses.

Spatial mismatch indicates that efficiency loss occurs when high-end brands detach from demand capacity. BOP theory emphasizes that standardized supply for low-income markets can effectively activate overlooked consumer potential. These two together constitute the fundamental mechanism through which RMI affects domestic demand.

In lower-tier markets, commercial infrastructure is weak. The main contradiction consumers face is the mismatch between the growing demand for quality and the backward local supply. An increase in RMI means a higher proportion of mass-market brands. This “downward compatibility” corrects spatial mismatch, producing a “Filling Effect” and driving retail sales. Therefore, the following is proposed:

H1: In lower-tier markets, RMI has a significant positive activation effect on urban domestic demand.

By contrast, the business ecosystem of mature markets has nearly reached saturation. Micro-adjustments in single formats are often zero-sum games, making the city insensitive to RMI changes. This shows an “elasticity desensitization” feature.

H2: In mature markets, RMI does not have a significant pull effect on urban domestic demand.

3. Research Design and Data Description

3.1 Data Sources and Sample Screening

To ensure the scientific nature of empirical research and the rigor of statistical inference, macro urban economic data, micro spatial data, and city level classifications are included. The specific sources of the core data are as follows:

(1) Macro economic data: Retail sales, GDP, and industrial structure data are manually matched from official statistical yearbooks.

(2) Micro spatial data (POI): Spatial data for “Starbucks” and “Chabaidao” are accurately extracted via the Gaode Map Open Platform with the authoritative domestic Narrow Gate Restaurant Brand Business Database. This reflects real commercial spatial structures.

(3) City grouping data: The City Commercial Charm Ranking is used as the measurement baseline. This list is currently the most widely recognized industry benchmark in China for the level of commercial infrastructure in cities.

The dataset was strictly cleaned. Cities with border changes, missing data, or too few stores were removed. Continuous macro variables were logarithmized. Ultimately, 150 observations of Strongly Balanced Panel Data were obtained. It laid a solid data foundation for subsequent causal inference.

3.2 Variables Selection and Measurement

Four types of indicators are established to comprehensively examine the impact of supply structure. The specific basis for selecting each variable and the measurement methods are shown in Table 1.

Table 1: Variable Definitions and Measurements

Variable Type	Variable	Symbol	Measurement
Dependent variable	Macro Consumption Vitality	Ln_sales_{it}	Ln (per capita retail sales)
Independent variable	Relative Mass-market Index	RMI_{it}	$RMI_{it} = \frac{N_{chabaidao,it}}{N_{starbucks,it} + N_{chabaidao,it}}$
Controls	Macro Economic Level	Ln_GDP_{it}	Ln (per capita GDP)
Controls	Industrial Structure	Ind_Struct_{it}	Tertiary industry GDP share (%)
Grouping Variable	City Commercial Level	$Grade_i$	Divided into mature markets and lower-tier markets

Dependent variable: Ln_sales_{it} is selected to measure consumption scale. The natural logarithm is executed to alleviate heteroscedasticity problem.

Independent variable: RMI_{it} is constructed from an internal structure fit perspective. “Starbucks” is chosen as the anchor for high premium thresholds. “Chabaidao” is chosen as the anchor for mass-market purchasing power. RMI ranges from 0 to 1. An index closer to 1 indicates the supply structure tends to align with mass purchasing power. Conversely, if it approaches 0, it indicates a tendency for high-end vacancies in the city's commercial supply.

Control variables: Ln_GDP_{it} controls the natural pull effect of basic wealth. Ind_Struct_{it} controls the interference of service sector development.

Grouping variable: $Grade_i$ divides the sample into mature markets and lower-tier markets based on the City Commercial Charm Ranking.

3.3 Model Specification

The specific econometric equation specification is as follows:

$$Ln\_sales_{it} = \beta_0 + \beta_1 RMI_{it} + \gamma_1 Ln\_GDP_{it} + \gamma_2 Ind\_Struct_{it} + \mu_t + \varepsilon_{it} \quad (1)$$

Where the subscript i and t represent the city and year, respectively; β_0 is the intercept term; μ_t is the year fixed effect; ε_{it} is the random error term.

RMI is a “Slow-moving Variable.” If City-specific fixed effects are forcibly controlled, valuable cross-sectional variance is irreversibly removed, causing the estimated coefficients to exhibit severe downward bias. Therefore, Pooled OLS is adopted. Cluster-Robust SE clustered at the city level is uniformly adopted to ensure the consistency and unbiasedness of statistical inference.

4. Empirical Results and Analysis

4.1 Descriptive Statistics

Before performing the estimation of measurement parameters, this paper conducts descriptive statistical analysis on various key macroeconomic and microeconomic variables (see Table 2).

Table 2: Descriptive Statistical Analysis of Major Variables

Variable	Obs	Max	Min	Mean	Std.Dev.
RMI_{it}	150	0.860	0.146	0.474	0.157
Ln_GDP_{it}	150	12.237	10.867	11.664	0.337
Ind_Struct_{it}	150	84.800	40.800	55.327	9.343
Ln_sales_{it}	150	9.063	7.674	8.386	0.302

From the statistical results in Table 2, it can be seen that Ln_sales_{it} shows a reasonable span in maximum and minimum values with a standard deviation of 0.302, indicating that there is indeed significant heterogeneity in the consumption base among different cities. The mean of RMI is 0.474, with a standard deviation of 0.157, providing rich cross-sectional variance to capture the “downward compatibility” effect.

4.2 Baseline and Heterogeneity Regression Analysis

According to the empirical estimation results in Table 3, this article provides the following analysis on the internal mechanism by which the optimization of commercial supply structure drives urban consumption vitality:

In lower-tier markets, the RMI estimation coefficient is 0.5353, highly significant at the 1% level. A 0.1-unit increase in RMI drives Ln_sales_{it} up by about 5.35%. H1 is verified.

In mature markets, the RMI coefficient is only 0.124 and fails the significance test ($P=0.547$). This asymmetric performance perfectly proves H2. High-tier cities are highly saturated; increasing mass-market proportions cannot produce substantial influence.

Ln_GDP_{it} is extremely significant in both models. It further confirms that the economic fundamentals and household income are the fundamental cornerstones driving retail prosperity. The R-squared values exceed 0.62. Cluster-Robust SE effectively eliminates endogenous interference from serial correlation.

Table 3: Analysis of Regression Results

Variable	Model 1 (Lower-tier markets)	Model 2 (Mature markets)
RMI_{it}	0.5353***	0.1238
Ln_GDP_{it}	0.7641***	0.8366***
Ind_Struct_{it}	0.0049	-0.0002
Year FE	YES	YES
Cluster	YES	YES
Obs	93	57
Constant	-1.0700	-1.3931
R^2	0.642	0.625

Note: *** represents significance at the 1% level.

4.3 Robustness and Endogeneity Tests

In the preceding baseline and heterogeneous group regressions, this paper confirms that RMI has a strong positive driving effect on urban macro consumption vitality in the sinking market. However, this empirical result may be subject to potential interference from endogeneity issues, particularly the challenge of reverse causality. Specifically, lower-tier cities with stronger overall consumption vitality and a more stable foundation of residents' purchasing power, may naturally attract more mass-market standardized restaurant chain brands, thereby passively raising the RMI index.

To overcome potential reverse causality, a one-period lagged independent variable is executed. Past supply structure determines current consumption ecosystems, but current consumption cannot reversely intervene in past brand locations, meeting the “cause-and-effect” chronological feature. The estimation results of the one-period lag robustness test are shown in Table 4:

Table 4: Robustness Test

Variable	Model
RMI_{it-1}	0.576***
Ln_GDP_{it}	0.7809***
Ind_Struct_{it}	0.006*
Year FE	YES
Cluster	YES
Obs	62
Constant	-1.339
R^2	0.654

Note: *** represents significance at the 1% level, * represents significance at the 10% level.

After excluding endogeneity, the lagged RMI coefficient reaches 0.576. The R-squared stabilizes at 0.654. It is strongly proven that the activation effect of “downward compatibility” remains stably present.

5. Conclusion

This paper empirically examines the impact of the physical commercial supply structure on urban consumption vitality. It is found that “downward compatibility” produces an extremely significant pull effect on lower-tier markets, effectively filling the supply-demand gap. Mature markets show obvious “system resilience” due to saturated facilities. Accordingly, it is suggested that low-tier cities should abandon blind high-end tendencies and increase support for high-quality mass-market formats. High-tier cities should focus on stock optimization and experiential retail, promoting the transformation of commercial supply from total expansion to refined structural adaptation.

There are still some shortcomings in this article. Specifically reflected in the following three aspects. Firstly, due to the difficulty of obtaining micro level store spatiotemporal big data, the RMI constructed in this paper only selected two benchmark freshly made beverage brands as measurement benchmarks. In the future, efforts can be made to expand the full category chain format to enrich the indicator dimensions; Secondly, the study

period of this article is from 2022 to 2024, which belongs to typical short panel data and fails to fully capture the dynamic time lag effects of the evolution of physical business networks under longer macroeconomic cycles; Finally, this article mainly focuses on the aggregation effect at the macro level of cities. In the future, it can further explore the spatial mismatch and micro mechanism of commercial resource distribution within cities by combining micro geographic grid data.

References

- [1] Couture, V., & Handbury, J. (2020). Urban revival in America. *Journal of Urban Economics*, 119, 103267.
- [2] Baviera-Puig, A., Buitrago-Vera, J., & Escriba-Perez, C. (2016). Geomarketing models in supermarket location strategies. *Journal of Business Economics and Management*, 17(6), 1205-1221.
- [3] Wang, T., Wang, Y., Zhao, X., & Fu, X. (2018). Spatial distribution pattern of the customer count and satisfaction of commercial facilities based on social network review data in Beijing, China. *Computers, Environment and Urban Systems*, 71, 88-97.
- [4] Kain, J. F. (1968). Housing segregation, negro employment, and metropolitan decentralization. *The quarterly journal of economics*, 82(2), 175-197.
- [5] SUN Bindong, PAN Xin, NING Yuemin. (2008). Analysis on Influence of Job-Housing Balance on Commute Travel in Shanghai. *Urban Planning Forum*, (1), 77-82.
- [6] Gobillon, L., Selod, H., & Zenou, Y. (2007). The Mechanisms of Spatial Mismatch. *Urban Studies*, 44(12), 2401-2427.
- [7] Nathan Schiff. (2015). Cities and product variety: evidence from restaurants, *Journal of Economic Geography*, 15 (6), 1085–1123.
- [8] Tang R, Hou G, Chen Y, et al. (2025). Spatial and temporal characteristics and driving mechanisms of recreational ecosystem services supply-demand mismatch in rapidly urbanizing areas: Evidence from the Yangtze River Delta. *Ecological Indicators*, 171, 113153.
- [9] Glaeser, E. L., Kim, H., & Luca, M. (2018). Nowcasting gentrification: Using Yelp data to quantify neighborhood change. *AEA Papers and Proceedings*, 108, 77-82.
- [10] London, T., & Hart, S. L. (2004). Reinventing strategies for emerging markets: beyond the transnational model. *Journal of international business studies*, 35(5), 350-370.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Extreme Weather Impacts on Stock Price Crash Risk: Evidence from China

Wenjing Ma and Jiusheng Chen*

School of Economics & Management, Chongqing Normal University, Chongqing, China

**Corresponding author: Jiusheng Chen.*

Abstract

Extreme weather events have increasingly imposed significant shocks on corporate operations and the financial system's stability. This study examines the impact of extreme weather on the stock price crash risk and its underlying mechanisms using a sample of A-share listed companies on the Shanghai and Shenzhen stock exchanges from 2010 to 2023. The results demonstrate that extreme weather significantly increases the stock price crash risk for listed companies. After addressing endogeneity issues and conducting a series of rigorous robustness tests, the core conclusions remain robust. Channel analyses reveal that extreme weather exacerbates stock price crash risk by deteriorating corporate fundamentals, reducing the quality of financial disclosures, and exacerbating information asymmetry. Heterogeneity analysis further indicates that the reinforcement effect is more pronounced among firms in the eastern region, firms with high-quality information disclosures, and non-state-owned enterprises. These findings contribute to the literature by identifying novel determinants of stock price crash risk and shedding light on the role of climate risk among investors.

Keywords

extreme climate, stock market crash risk, information asymmetry, information disclosure

1. Introduction

Against the backdrop of profound changes in the global climate governance system and the incorporation of “dual carbon” targets into national development strategies, climate change has evolved from a long-term environmental issue into an urgent factor affecting economic and social systems. The Sixth Assessment Report of the United Nations Intergovernmental Panel on Climate Change [1] clearly states that the frequency, intensity, and duration of extreme weather events have increased substantially worldwide, imposing systematic shocks on regional production networks and firm operations. Over the past four decades, European countries have suffered losses ranging from 450 billion to 520 billion euros; China has also been severely affected by climate-related disasters, with the 2008 nationwide cold snap, the 2020 heavy rainfall in the Yangtze and Huai River basins, and the 2021 “extreme rainstorm” in Henan causing direct economic losses of 151.7 billion, 132.2 billion, and 120.1 billion yuan, respectively. Such shocks not only cause direct damage to physical assets and supply chains but also transmit potential risks to financial markets through channels such as corporate financial stability, disclosure practices, and investor expectations.

Existing research on the risk of stock price crashes is primarily based on theories of information asymmetry and agency conflicts [2], which suggest that management engages in hoarding bad news. This is primarily manifested in opaque financial reporting [3] and poor quality of information disclosure [4]. These behaviors are regarded as key channels amplifying the stock price crash risks. Within this framework, extreme weather events may exacerbate this risk through two channels. The first is the physical risk channel that extreme weather events directly lead to corporate asset impairments [5], production disruptions, and cash flow volatility [6], thereby deteriorating corporate fundamentals and increasing operational uncertainty. The second is the transition risk and disclosure channel. Under increasingly stringent climate regulations, firms may face regulatory penalties and reputational damage due to inadequate environmental disclosures or “greenwashing,” further triggering a crisis of market confidence [7]. Both channels may strengthen managers’ incentives to conceal bad news, ultimately causing a stock price crash when negative information is released in a concentrated manner.

However, existing literature predominantly focuses on the impact of macro-level climate policies on corporate environmental investments or green innovation, finding that climate policy uncertainty crowds out both real-economy investments and green innovation investments [8] and significantly increases the volatility of the S&P Green Bond Index, the Global Clean Energy Index, and the Dow Jones Sustainability Index [9]. At the same time, environmental initiatives such as green product innovation and pollution control are often accompanied by severe information asymmetry and high uncertainty, which may have a significant negative impact on financial performance [10]. In addition, major adjustments in climate policy can also amplify the green credit bubble effect, significantly increasing banks’ risk levels and exacerbating the climate transition risks they face [11]. Corporate uncertainty is the underlying mechanism driving stock price volatility [12]. The core logic lies in the fact that uncertainty affects future cash flow and capital cost assessments, thereby amplifying stock price volatility [3]. The uncertainties triggered by extreme weather do not operate in isolation. On the contrary, they are intertwined with uncertainty in the macroeconomic, financial, and energy policy domains and reinforce one another. This interaction creates a coupling effect that significantly increases stock market risk.

Against this backdrop, investigating the impact of extreme weather on the risk of stock price crashes is particularly significant. Extreme weather events signify not only a direct exogenous shock but also represent a highly complex climate risk exposure at the corporate level. This directly influences investors’ assessments of a company’s asset value, supply chain stability, and long-term sustainable operational capacity [5], thereby disrupting the underlying assumptions of cash flow estimates [6]. More importantly, extreme climate risks are deeply embedded within a broad network of exogenous uncertainties, particularly those arising from global climate policy trajectories, the evolution of regulatory frameworks, and the restructuring of technological pathways^[8, 9]. Such exogenous uncertainties will trigger a fundamental restructuring of economic rules, exerting profound and destabilizing effects on corporate fundamentals through channels such as direct asset impairments, disruptions to production operations, shifts in industry competitive landscapes, and carbon trading costs. In this process, management’s incentive to conceal adverse news is significantly strengthened in order to mask the deterioration in fundamentals caused by climate shocks. Once this accumulated concealed bad news is eventually disclosed, the market may experience a sudden sell-off, thereby triggering the risk of a stock market crash.

Hence, using data of A-share firms listed on the Shanghai and Shenzhen stock exchanges from 2010 to 2023, this study systematically investigates the impact of extreme weather on stock price crash risk, its transmission mechanisms, and its heterogeneity. The findings reveal that extreme weather significantly increases the stock price crash risk. The conclusion that remains robust after a series of robustness tests, including replacing the dependent variable, adding control variables, and excluding 2020 data. Heterogeneity analyses show that the effect is more pronounced among firms in the eastern region, firms with high disclosure levels, and non-state-owned enterprises. Mechanism tests indicate that extreme temperatures exacerbate stock price crash risk mainly by intensifying information asymmetry and deteriorating corporate fundamentals. This study not only extends the theoretical boundaries of climate finance in emerging markets but also provides empirical evidence to assist regulators in refining climate disclosure frameworks and to help investors identify and manage climate-related risks.

The marginal contributions of this paper are primarily reflected in the following three aspects. First, we contribute to the emerging literature on the determinants of stock price crash risk. Existing literature has

identified factors influencing stock price crash risk from various perspectives, such as information disclosure [13], managerial characteristics [14], corporate policies and internal governance [15], and culture [16], but none have considered the impact of climate factors. This study finds that extreme weather also increases the risk of stock price crashes, expanding our understanding of the specific determinants of this risk.

Second, most existing studies have merely established a macro-level correlation between extreme weather and market volatility, while remaining notably deficient in empirically examining the micro-level transmission mechanisms. This study is the first to empirically separate and validate two parallel and complementary mediating channels, information asymmetry and corporate fundamentals, within a unified framework. By tracing the complete causal chain through which macro-level climate shocks influence micro-level firm behavior and ultimately shape financial market outcomes, this study provides crucial micro-level evidence from China regarding the financial transmission of climate risks, thereby deepening the understanding in this field.

2. Literature Review

2.1 Stock Price Crash Risks and Their Determinants

Stock price crash risk refers to the probability that a specific stock's intrinsic value will plummet to an extremely negative level, typically due to unexpected events, the release of negative internal corporate news, or a sudden shift in investor sentiment. Regarding the origins of stock price crash risk, research suggests that information asymmetry between companies and investors, coupled with management's deliberate concealment of negative news, is the primary cause [17]. This theory stems from the widespread principal-agent problem within firms, positing that corporate management will protect its own interests by concealing negative information about the company. Behind this behavior, the company is likely facing deteriorating fundamentals such as insufficient innovation, a lack of market competitiveness in its products or services, profit pressures, and tight cash flow [18]. To prevent this negative information from reaching the capital markets, which could trigger declining investor expectations and other contagion effects, management is forced to conceal the bad news. However, a company cannot hide this negative information indefinitely; when the accumulated negative information reaches a level that the company cannot bear, it will gradually be exposed, shocking the market, increasing stock price volatility, and ultimately leading to a stock price crash [19]. Therefore, a stock price crash is not only a direct consequence of information asymmetry and management's concealment of adverse news, but also a retrospective punishment by the market for the deterioration of a company's fundamentals and the behavior of "poor-quality companies".

From the perspective of internal factors, the relevant literature primarily examines corporate governance, information disclosure, and management characteristics. Yuan and Shao [20] found that corporate governance structure and the level of financial information disclosure have a significant impact on the probability of a stock price crash; the less adequate the information disclosure, the higher the probability of a stock price crash. Liu and Li [21], using data from A-share listed companies in 2011, found that firms can reduce the likelihood of a stock price crash by increasing the disclosure of internal control information, and this effect is more pronounced in firms with higher levels of information asymmetry and weaker profitability. An empirical study by Zhang X et al. [22], utilizing a large language model, indicates that the degree of readability is strongly negatively linked to firm stock price crash risk. From the perspective of managerial behavior, Fu F et al. [23] found that long-term investments by institutions significantly mitigate risk by curbing managerial myopia, enhancing transparency, and improving accountability, thus deterring self-serving behavior. Pu et al. [24] found that board faultlines influence corporate risk-taking by disrupting internal board order, lowering its monitoring willingness and ability, and reducing investment efficiency. In summary, regardless of the specific factors involved, enhancing corporate transparency and reducing management's concealment of information are key to mitigating the risk of stock price crashes.

In addition, macroeconomic policies and the broader environment can influence corporate behavior and, consequently, the risk of stock price crashes. Dai Bingbin et al. [25] found in their study of monetary policy that tight monetary policy is positively correlated with the risk of stock price crashes, and that insufficient liquidity in funds and individual stocks significantly amplifies this effect. Green credit policies increase the risk of stock price crashes for heavily polluting firms by imposing financial constraints and reducing the quality

of information disclosure [26]. Tax authority monitoring improves information disclosure quality through tax compliance channels, and the effect is more significant in firms with weak external supervision, weak internal governance, and firms located in non-Eastern areas and areas with lower fiscal pressure [27].

2.2 Economic Consequences of Extreme Climate Change

As a key component of climate risk, physical climate risks have received significant attention in recent years. According to the literature in this field, climate risks can be categorized into “physical risks” and “transition risks” [28]. “Physical risks” refer to direct economic and financial losses caused by climate risks. Extreme weather events, such as storms, floods, and heatwaves, significantly reduce corporate productivity, alter corporate financing structures, and have direct negative impacts on the macro economy, leading to disruptions in economic activity and a decline in the value of corporate assets. Consequently, the returns on their financial assets also drop sharply [29], thereby reducing corporate profitability [30].

At the same time, such risks can exacerbate volatility in financial markets, further triggering financial risks. This is because, when investors are paying close attention to climate change, the accumulation of physical risks can easily trigger panic among investors, which then spreads to financial markets such as the stock market [31]. When risk events such as abnormal heat waves and extreme droughts impact businesses and markets, investors tend to exhibit pronounced overreactions, thereby significantly depressing bond and stock prices [32]. Physical risks arising from extreme weather are also linked to insurance mechanisms for climate disasters. The occurrence of extreme weather may lead to impairment of fixed assets, substantially increase insurance claims, exacerbate volatility in the insurance sector, and further transmit these effects to other financial institutions, such as real estate firms [33]. Addoum [34] found that extreme temperatures affect the profitability of over 40% of industries, with temperature-sensitive sectors being particularly vulnerable. Consequently, central banks worldwide have issued warnings regarding the potential impact of climate risks on financial stability. Policymakers have also emphasized that assessing climate change as a potential source of systemic risk is crucial for ensuring the stability of the financial system [35].

3. Theoretical Mechanisms and Research Hypotheses

3.1 Theoretical Framework of How Extreme Climate Shocks Affect Stock Price Crash Risk

The transmission of extreme climate events to the risk of corporate stock price crashes is essentially a process of transforming physical risks into financial risks. At its core lies how climate shocks disrupt the stability of corporate operations and ultimately trigger nonlinear crash effects in capital markets. Extreme climate events can directly damage a company’s tangible assets, such as factories and equipment, disrupt production processes, and lead to reduced production capacity and revenue [34]. For example, floods may submerge manufacturing production lines, extreme heat may reduce agricultural output, and hurricanes may destroy energy companies’ infrastructure. The destruction of physical assets caused by disasters such as hurricanes and floods forces firms to bear high restoration costs; such expenditures are often urgent and cannot be deferred, significantly depleting firms’ cash reserves [36]. Furthermore, the decline in the collateral value of damaged assets triggers risk reassessments by financial institutions, leaving firms facing difficulties in extending debt maturities and a sharp rise in financing costs [37]. When these dual financial pressures reach a critical point, a firm’s debt-repayment capacity will deteriorate substantially; at this stage, any signal of debt default could serve as the trigger for a precipitous plunge in stock prices.

The strong visual impact of extreme weather events and their widespread media coverage significantly alter collective perceptions in capital markets. Environmental anxiety triggered by disaster footage can easily lead to irrational panic, causing investors to indiscriminately sell off shares of climate-vulnerable companies and even entire industry sectors [32]. Behavioral resonance among institutional investors further amplifies this effect: ESG funds, driven by compliance requirements, passively reduce their holdings of impacted companies’ stocks, potentially triggering a liquidity spiral; meanwhile, short sellers seize the opportunity to concentrate their positions on firms exposed to climate risks, accelerating the divergence of prices from fundamentals. This attention-driven trading behavior essentially amplifies the tail-risk attributes of climate shocks, causing the distribution of stock returns to exhibit significant negative skewness. This mechanism demonstrates that climate disasters not only affect the intrinsic value of firms but also create self-fulfilling conditions for stock

price crashes by reshaping market expectations. Based on the above analysis, this paper proposes the following research hypothesis:

H1: Extreme temperatures exacerbate the risk of stock price crashes.

3.2 The Mechanism Through Which Extreme Climate Shocks Influence the Risk of Stock Price Collapse

The distinctiveness of climate risk lies in its difficult-to-quantify uncertainty; this ambiguity exacerbates information asymmetry. Based on agency theory, management may deliberately conceal negative information resulting from extreme climate events to protect their reputation, compensation, or job security [4]. Companies in high-carbon-emission industries may engage in “green washing” to embellish their climate response measures, deliberately concealing their environmental risk exposures [38]. Faced with long-term climate threats, corporate management’s decision-making may fall into the trap of “defensive myopia.” To maintain short-term performance, executives often cut investments in disaster-prevention infrastructure or accelerate the exercise and sale of stock options to lock in personal gains. At the board level, a lack of climate expertise often leads to an underestimation of the severity of physical risk exposure, resulting in the absence of climate audit mechanisms. This governance failure causes companies to gradually lose their ability to adapt to risks, meaning even minor climate shocks can result in disproportionate losses. However, concealment leads to the continuous accumulation of negative information. When the costs of concealment exceed the benefits, a flood of negative information can suddenly hit the market, triggering a stock price crash [39].

Given the sudden nature and uncertainty regarding the severity of extreme weather events, market reactions suggest that a deterioration in corporate fundamentals lowers investors’ expectations of a company’s future value. However, if negative information is not disclosed in a timely and adequate manner, investors may maintain an erroneous valuation of the company due to information lag. For example, asset impairment losses caused by extreme weather may be recognized late, or the impact of supply chain disruptions may be underestimated, providing management with an excuse to withhold information. They may attribute declining performance to “force majeure” rather than their own inadequate risk management, thereby reducing market accountability. Based on the above analysis, this paper proposes the following research hypothesis:

H2: Extreme temperatures exacerbate the risk of stock price crashes by deteriorating corporate fundamentals and increasing the degree of information asymmetry.

4. Data and Methods

4.1 Data Sources and Processing

We selected Chinese A-share listed companies on the Shanghai and Shenzhen stock exchanges from 2010 to 2023 as the study sample and performed the following screening and processing: (1) Excluded companies in the financial sector; (2) Excluded companies listed for less than one year, those that had been delisted, or those with suspended listings; (3) Removed missing values; (4) To eliminate the influence of outliers, the values of continuous variables at the 1st and 99th percentiles were truncated. After the above data screening and processing, a total of 24,148 sample data points were obtained. The listed company data used in this study were sourced from the CSMAR database.

4.2 Variable Definitions and Measurement

4.2.1 Dependent Variables

Following the relevant research by Wang et al. [40], this study selects the negative return skewness coefficient (NCSKEW) and the down-up volatility ratio (DUVOL) as metrics for stock price crash risk. A higher value of NCSKEW indicates a greater degree of negative weekly stock returns for the company and a higher risk of stock price collapse; a higher value of DUVOL indicates a greater amplitude of stock price fluctuations around the mean, implying a higher risk of stock price collapse. The specific calculation formulas are as follows:

$$NCSKEW_{i,t} = -\frac{n(n-1)^{3/2} \sum w_{i,t}^3}{(n-1)(n-2)(\sum w_{i,t}^2)^{3/2}} \quad (1)$$

$$DUVOL_{i,t} = \log \left\{ \left[\frac{(n_u-1) \sum_{Down} w_{i,t}^2}{(n_d-1) \sum_{up} w_{i,t}^2} \right] \right\} \quad (2)$$

Since the relevant literature has already provided a detailed introduction to the calculation methods and procedures for the above two variables, this paper will not elaborate further.

4.2.2 Core Explanatory Variables

This paper uses extreme high temperatures (HTD) and extreme low temperatures (LTD) as representatives of extreme climate for empirical analysis. These two types of extreme climate events often lead to common meteorological disasters. Extreme heat can trigger heatwaves, causing shortages in water and electricity supply and reducing labor productivity; extreme cold causes frost and cold damage, resulting in severe losses to agricultural production, transportation, and communications.

The data is sourced from the Global Climate Risk Integration Database (GCRID). The period from 1973 to 1992 was selected as the climate baseline. Using daily observational data from meteorological stations during the baseline period and the percentile relative threshold method, extreme thresholds for the city's meteorological indicators were defined, and the number of days with extreme temperature events was calculated for each year in the sample period from 1993 to 2023. Specifically, the daily maximum and minimum temperature records for the same date in each year of the climate reference period (e.g., January 1 of each year), sorted in ascending order, and the 90th percentile and 10th percentile values as the extreme high-temperature threshold and extreme low-temperature threshold, respectively. When the maximum temperature in a city on a given day exceeds the extreme high-temperature threshold for that day, it is considered that an extreme high-temperature event occurred in that city on that day. Finally, the daily-level data are aggregated to obtain the number of days with extreme climate events at the annual level. The number of days with extreme low-temperature events is extracted using the same method.

4.2.3 Control Variables

To enhance the reliability of the research findings, this study controls for relevant factors that may influence the risk of stock price crashes in listed companies. Drawing on the research by Kim [41] and Chang [42], we use firm-level variables such as firm size (SIZE), return on assets (ROA), leverage ratio (LEV), monthly excess turnover (Dturn), property, plant, and equipment (PPE), and Tobin's Q to control for the impact of fundamental firm characteristics on the risk of stock price crashes; we select macroeconomic variables (GDP) to control for the impact of macroeconomic characteristics on stock price crashes. See Table 1 for specific variable definitions.

Table 1: Definitions and Descriptions of Main Variables

Variable Type	Variable Name	Variable Symbol	Variable Definition
Dependent Variable	Negative Return Skewness	NCSKEW	See formula (1)
	Stock Price Volatility	DUVOL	See formula (2)
Independent Variable	Extreme High Temperature	HTD	Number of days exceeding the extreme high temperature threshold
	Extreme Low Temperature	LTD	Number of days exceeding the extreme low temperature threshold
Control Variable	Firm Size	Size	Logarithm of total assets
	Leverage	Lev	Ratio of total liabilities to total assets
	Return on Assets	ROA	Ratio of after-tax net profit to total assets
	Fixed Assets	PPE	Ratio of fixed assets to total assets
	Excess Turnover Rate	Dturn	Average daily turnover rate within the year
	Tobin's Q	Tobinq	Ratio of market value to total assets
	Regional Economic Development	GDP	Natural logarithm of GDP in the firm's region

4.3 Empirical Model Design

We construct the following empirical model to estimate the impact of climate risk on stock price crash risk:

$$Crash_{i,t} = \alpha_0 + \alpha_1 temperature_{i,t} + \alpha_2 Controls_{i,t} + \mu_i + \gamma_i + \varepsilon_{i,t} \quad (3)$$

The dependent variable is firm-specific stock price crash risk (Crash), measured by the negative return skewness (NCSKEW) and the down-to-up volatility ratio (DUVOL), respectively. The independent variable is extreme temperature, measured by extreme high temperature (HTD) and extreme low temperature (LTD), respectively. Controls represent the corresponding control variables. μ_i and γ_i denote firm and time fixed effects, respectively, and $\varepsilon_{i,t}$ is the error term. If the research hypothesis holds, the estimated coefficient α_1 should be significantly positive.

5. Analysis of Empirical Results

5.1 Descriptive Statistics

Table 2 reports the main descriptive statistics. In the sample, the mean of the skewness coefficient for negative stock returns is -0.427, with a standard deviation of 0.717; the minimum and maximum values are -2.670 and 1.614, respectively; the mean of stock price volatility is -0.291, the standard deviation is 0.475, and the minimum and maximum values are -2.932 and 3.375, respectively. The above indicates significant differences in the risk of stock price crashes across companies, which is consistent with existing research. The mean for extreme low temperatures is 17.263, with a standard deviation of 8.047; the minimum and maximum values are 2.703 and 39.865, respectively. The mean for extreme high temperatures is 52.081, with a standard deviation of 14.2992; the minimum and maximum values are 0.000 and 111.712, respectively. This indicates that there are also significant differences among the various extreme climate indicators, which provides a research scope for the empirical analysis in this paper.

Table 2: Descriptive Statistics

Variable	Sample Size	Mean	Std	Min	Median	Max	Skewness	Kurtosis
NCSKEW	24148	-0.427	0.717	-2.670	-0.374	1.614	-0.364	3.956
DUVOL	24148	-0.291	0.475	-2.932	-0.288	3.375	0.081	3.814
LTD	24148	17.263	8.047	2.703	16.216	39.865	0.464	2.635
HTD	24148	52.081	14.292	0.000	51.351	111.712	0.062	3.838
Lev	24148	0.434	0.205	0.051	0.427	0.908	0.176	2.198
PPE	24148	20.117	16.021	0.148	16.456	69.621	1.034	3.611
ROA	24148	3.460	6.018	-26.214	3.510	18.720	-1.416	9.000
Size	24148	22.376	1.323	19.916	22.169	26.465	0.793	3.529
Dturn	24148	1.683	1.252	0.178	1.330	6.470	1.506	5.396
Tobinq	24148	1.985	1.223	0.833	1.606	8.632	2.592	11.697
GDP	24148	11.563	0.495	10.050	11.650	12.223	-0.906	3.294

5.2 Baseline Regression

Table 3 presents the baseline regression results evaluating the impact of extreme cold and extreme heat on the risk of stock price crashes. The results of the univariate regression in columns (1) and (5) indicate that, at the 1% significance level, the LTD coefficient is positive and statistically significant. Columns (2) and (6) report the results after controlling for covariates; the LTD coefficients are 0.0041 and 0.0024, respectively, both of which are significant at the 1% level. These empirical results strongly support Hypothesis 1, confirming a positive correlation between extremely low temperatures and the risk of stock price crashes; similarly, the regression results for extreme high temperatures in columns (2) and (7) indicate a positive correlation with the risk of stock price crashes at the 1% significance level, and the results remain significant after controlling for variables.

Meanwhile, the coefficient for firm size in Table 3 is significantly positive at the 1% level, indicating that larger firms face a higher risk of stock price crashes. One possible explanation is that larger firms have more complex business structures and broader market influence, making them more susceptible to various uncertainties [43].

Table 3: Baseline regression results

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	NCSKEW				DUVOL			
ltd	0.0050*** (7.5397)	0.0041*** (5.4301)			0.0031*** (7.0611)	0.0024*** (4.7924)		
htd			0.0027*** (6.3568)	0.0021*** (4.7023)			0.0021*** (7.1284)	0.0014*** (4.7087)
lev		-0.0819 (-1.5227)		-0.0691 (-1.2875)		-0.0593 (-1.6439)		-0.0522 (-1.4488)
ppe		-0.0014** (-1.9788)		-0.0015** (-2.0070)		-0.0012*** (-2.5767)		-0.0013*** (-2.6036)
roa		-0.0020* (-1.7201)		-0.0017 (-1.4366)		-0.0027*** (-3.5174)		-0.0025*** (-3.2560)
size		0.0621*** (5.0687)		0.0442*** (3.6680)		0.0296*** (3.5914)		0.0187** (2.3091)
dturn		-0.0879*** (-17.3189)		-0.0902*** (-17.9429)		-0.0621*** (-18.3488)		-0.0633*** (-18.8350)
tobinq		0.0237*** (4.0580)		0.0187*** (3.2450)		0.0052 (1.2704)		0.0023 (0.5763)
gdp		-0.0207 (-0.7709)		-0.0644** (-2.4087)		0.0202 (1.1362)		-0.0071 (-0.4040)
cons	-0.5136*** (-41.1097)	-1.4743*** (-5.2465)	-0.5687*** (-25.0758)	-0.6002** (-2.3561)	-0.3449*** (-41.4502)	-1.0728*** (-5.6774)	-0.3979*** (-25.9786)	-0.5422*** (-3.1846)
N	24148	24148	24148	24148	24148	24148	24148	24148
R ²	0.1662	0.1807	0.1656	0.1804	0.1599	0.1774	0.1601	0.1774
F	56.8467	51.3714	40.4089	50.0159	49.8590	58.8419	50.8134	58.5751

5.3 Robustness Tests

The baseline regression results indicate that extreme temperatures are associated with an increased risk of stock market crashes. Although control variables have been included, potential concerns regarding endogeneity and omissions bias may still undermine the validity of these findings. To address these concerns and enhance the robustness of the results, the following methods were employed to conduct robustness and endogeneity tests.

5.3.1 Excluding 2020

The COVID-19 pandemic triggered severe volatility in global stock markets, presenting financial markets with unprecedented challenges. During this period, an increase in insider trading exacerbated information asymmetry in the market, which in turn intensified irrational investor sentiment and ultimately led to sharp declines in stock prices [44]. To isolate the impact of specific events on the stock market, we conducted a robustness test by excluding the 2020 samples and re-estimating the baseline regression model. Table 4 presents the regression results. Compared to the baseline regression, the coefficient between extreme temperatures and the risk of stock price crashes has decreased, but the overall significance and robustness of the results remain unaffected.

5.3.2 Replacing the explanatory variable with CCPRI

We replace the explanatory variable “extreme temperatures” with the Climate Physical Risk Index (CCPRI) to examine its relationship with the risk of stock price crashes. The results in columns (5) and (6) of Table 4 show that the Climate Physical Risk Index has a significant positive correlation with the risk of stock price crashes at the 1% level, confirming the robustness of the baseline regression.

5.3.3 Including control variables

In addition to the explanatory variables in the baseline regression, several other factors may influence the risk of stock price crashes. Existing research indicates that analyst coverage [45] and financial constraints [46] play a key role in influencing the risk of stock price crashes, while stronger corporate governance measures significantly reduce a firm’s financial risk [47]. Therefore, we introduce additional control variables to address the potential impact of omitted variable bias. Specifically, the omitted variables include financial constraints (SA) and control variables related to internal and external corporate governance: board size (board) and the proportion of independent directors (Indep). Table 5 presents the impact of extreme weather on stock crash risk after incorporating these additional control variables into the baseline. Consistent with the findings in

Table 2, the coefficients for LTD and HTD remain positive and statistically significant, reaffirming the role of climate risk in increasing the likelihood of stock price crashes.

Table 4: Robustness Tests

	Excluding 2020				Replace with CCPRI	
	(1)	(2)	(3)	(4)	(5)	(6)
	NCSKEW	DUVOL	NCSKEW	DUVOL	NCSKEW	DUVOL
ltd	0.0038*** (4.8188)	0.0021*** (3.9787)				
htd			0.0020*** (4.1740)	0.0013*** (4.0798)		
CCPRI					0.0038*** (4.4926)	0.0023*** (4.0328)
lev	-0.0880 (-1.5662)	-0.0663* (-1.7686)	-0.0769 (-1.3710)	-0.0604 (-1.6169)	-0.0724 (-1.3477)	-0.0539 (-1.4965)
ppe	-0.0015* (-1.9408)	-0.0012** (-2.4722)	-0.0015** (-1.9788)	-0.0013** (-2.5025)	-0.0014* (-1.9239)	-0.0012** (-2.5258)
roa	-0.0022* (-1.7880)	-0.0027*** (-3.2419)	-0.0019 (-1.5325)	-0.0025*** (-3.0152)	-0.0018 (-1.5666)	-0.0026*** (-3.3838)
size	0.0688*** (3.8880)	0.0321*** (3.7488)	0.0529*** (4.1855)	0.0228*** (2.7007)	0.0497*** (4.1433)	0.0224*** (2.7726)
dturn	-0.0840*** (-15.6786)	-0.0590*** (-16.6383)	-0.0858*** (-16.1432)	-0.0598*** (-16.9638)	-0.0897*** (-17.8230)	-0.0631*** (-18.7621)
tobinq	0.0249*** (3.9372)	0.0050 (1.1362)	0.0196*** (3.1577)	0.0022 (0.5059)	0.0195*** (3.3851)	0.0027 (0.6823)
gdp	-0.0228 (-0.8139)	0.0220 (1.1969)	-0.0655** (-2.3559)	-0.0033 (-0.1821)	-0.0461* (-1.7435)	0.0053 (0.3025)
cons	-1.5995*** (-5.5432)	-1.1445*** (-5.8987)	-0.7749*** (-2.9518)	-0.6701*** (-3.8303)	-0.9339*** (-3.6983)	-0.7590*** (-4.4837)
N	22007	22007	22007	22007	24148	24148
r2	0.1908	0.1865	0.1905	0.1865	0.1803	0.1771
F	42.9708	49.1117	42.1025	49.2123	49.5701	57.7199

Table 5: Robustness tests with including control variables

	(1)	(2)	(3)	(4)
	NCSKEW	DUVOL	NCSKEW	DUVOL
ltd	0.0040*** (5.2165)	0.0024*** (4.7064)		
htd			0.0024*** (5.2462)	0.0016*** (5.1118)
sa	0.1401** (2.4122)	0.0641* (1.6514)	0.2214*** (3.7925)	0.1155*** (2.9552)
indep	0.0048*** (3.0767)	0.0029*** (2.8414)	0.0049*** (3.1581)	0.0029*** (2.9173)
board	-0.0045 (-0.6979)	-0.0051 (-1.2089)	-0.0034 (-0.5336)	-0.0045 (-1.0605)
lev	-0.0820 (-1.5246)	-0.0596* (-1.6528)	-0.0723 (-1.3478)	-0.0542 (-1.5050)
ppe	-0.0015** (-2.1069)	-0.0013*** (-2.6773)	-0.0016** (-2.1804)	-0.0013*** (-2.7473)
roa	-0.0022* (-1.8958)	-0.0027*** (-3.5987)	-0.0020* (-1.7537)	-0.0026*** (-3.4708)
size	0.0713*** (5.5436)	0.0342*** (3.9722)	0.0588*** (4.6326)	0.0266*** (3.1350)
dturn	-0.0868*** (-16.9679)	-0.0617*** (-18.0296)	-0.0877*** (-17.2539)	-0.0621*** (-18.2118)
tobinq	0.0228*** (3.8922)	0.0047 (1.1636)	0.0177*** (3.0756)	0.0018 (0.4402)
gdp	0.0242 (0.7126)	0.0384* (1.6827)	0.0104 (0.3061)	0.0299 (1.3163)
cons	-1.7963*** (-5.7023)	-1.2022*** (-5.6701)	-1.1046*** (-3.7425)	-0.7804*** (-3.9412)
N	24135	24135	24135	24135
R ²	0.1816	0.1782	0.1816	0.1784
F	38.8511	44.2486	38.7004	44.6506

5.4 Mediation Effect

The amplifying effect of climate risk on the risk of stock price crashes was verified in the preceding section. How, then, does climate risk increase the risk of stock price crashes? Based on the theoretical analysis presented earlier, this section examines the mechanism from two dimensions, the information asymmetry and corporate fundamentals, to reveal the underlying mechanism through which extreme temperatures affect the risk of stock price crashes.

5.4.1 Information Asymmetry

Following the approach of Hutton et al. (2009), we utilize the manipulable accruals indicator (DisAcc) calculated using the Modified Jones Model to examine the impact of extreme temperatures on the firm's information environment from the perspective of information quality. A higher value of manipulable accruals indicates lower quality of financial disclosure and a higher degree of information asymmetry. The test results are shown in Table 10. Columns (1) and (4) present the regression results for extreme low temperatures and extreme high temperatures on manipulable accruals, both showing a significant positive relationship at the 1% level. This indicates that extreme temperatures affect the quality of corporate financial disclosures and exacerbate information asymmetry. Columns (2) and (3), as well as (5) and (6) present the regression results after including mediating variables; the impact of extreme temperatures on the risk of stock price crashes remains significantly positive at the 1% level. This indicates that information asymmetry serves as a mediating variable in the relationship between extreme temperatures and the risk of stock price crashes, meaning that extreme temperatures exacerbate the risk of stock price crashes by influencing the degree of information asymmetry.

Table 6: Testing for Mediating Effects—Information Asymmetry

	(1)	(2)	(3)	(4)	(5)	(6)
	DisAcc	NCSKEW	DUVOL	DisAcc	NCSKEW	DUVOL
ltd	0.0002**	0.0041***	0.0024***			
	(2.5615)	(5.3708)	(4.7263)			
htd				0.0002***	0.0021***	0.0014***
				(4.3974)	(4.5919)	(4.5880)
DisAcc		0.2197***	0.1595***		0.2172***	0.1574***
		(3.9488)	(4.2351)		(3.8946)	(4.1714)
lev	0.0064	-0.0833	-0.0603*	0.0069	-0.0706	-0.0532
	(0.7142)	(-1.5505)	(-1.6737)	(0.7659)	(-1.3168)	(-1.4802)
ppe	-0.0001	-0.0014*	-0.0012**	-0.0001	-0.0014**	-0.0012**
	(-0.9828)	(-1.9511)	(-2.5452)	(-1.0013)	(-1.9791)	(-2.5720)
roa	0.0071***	-0.0036***	-0.0038***	0.0072***	-0.0032***	-0.0036***
	(40.3501)	(-2.9160)	(-4.7234)	(40.3928)	(-2.6361)	(-4.4619)
size	-0.0019	0.0625***	0.0299***	-0.0031*	0.0448***	0.0192**
	(-1.1220)	(5.1050)	(3.6290)	(-1.8409)	(3.7248)	(2.3692)
dturn	-0.0007	-0.0877***	-0.0619***	-0.0008	-0.0900***	-0.0631***
	(-1.1962)	(-17.2923)	(-18.3171)	(-1.2531)	(-17.9141)	(-18.8013)
tobinq	-0.0028***	0.0243***	0.0056	-0.0030***	0.0193***	0.0028
	(-3.3591)	(4.1667)	(1.3816)	(-3.6237)	(3.3623)	(0.6953)
gdp	-0.0336***	-0.0134	0.0255	-0.0369***	-0.0564**	-0.0013
	(-9.3738)	(-0.4950)	(1.4347)	(-10.0802)	(-2.1011)	(-0.0741)
cons	0.4078***	-1.5639***	-1.1378***	0.4646***	-0.7011***	-0.6153***
	(11.6608)	(-5.5409)	(-6.0004)	(13.5537)	(-2.7330)	(-3.5904)
N	24148	24148	24148	24148	24148	24148
r2	0.3012	0.1812	0.1781	0.3016	0.1809	0.1780
F	313.1374	47.3298	54.2672	313.5195	46.1835	54.0392

5.4.2 Deteriorate the Firm's Fundamentals

Previous research has shown that a deterioration in a company's fundamental performance increases the likelihood that management will conceal bad news, which is a major factor contributing to the risk of a stock price collapse. Furthermore, extreme weather directly impacts a company's revenue and assets, thereby undermining its profitability and financial performance.

Evidence provided by Rao [48] indicates that following extreme rainfall, the market capitalization of rainfall-sensitive companies declines significantly. Managers have an incentive to manipulate earnings to conceal poor performance caused by extreme weather events. Stock prices are typically positively correlated with corporate earnings and profitability; therefore, deteriorating performance amplifies the risk of stock price declines and may even lead to a stock price crash. If this view holds, the relationship between extreme weather and corporate fundamentals is expected to be negative. Following previous literature [49], we use Altman's Z-SCORE as a proxy for corporate fundamentals. A higher Z-SCORE value indicates a lower probability of default, suggesting that the company has stronger fundamentals.

The test results are shown in the figure. Columns (1) and (4) present the regression results for the effects of extreme low temperatures and extreme high temperatures on the Z-SCORE, respectively. These results are significantly negative at the 1% and 5% levels, indicating that extreme temperatures may reduce corporate financial resilience through operational shocks, cash flow volatility, and financing constraints, thereby deteriorating corporate fundamentals. Columns (2) and (3), as well as columns (5) and (6), present the regression results after controlling for mediating variables. The impact of extreme temperatures on the risk of stock price crashes remains significantly positive at the 1% level, while the coefficient of Z-SCORE on the risk of stock price crashes is significantly negative at both the 1% and 10% levels. This suggests that extreme weather, by weakening Z-SCORE, indirectly amplifies the risk of stock price crashes by weakening ZSCORE and deteriorating corporate fundamentals..

Table 7: Mediating effects test based on deterioration in company fundamentals

	(1)	(2)	(3)	(4)	(5)	(6)
	ZSCORE	NCSKEW	DUVOL	ZSCORE	NCSKEW	DUVOL
ltd	-0.0088*** (-2.6920)	0.0041*** (5.3745)	0.0024*** (4.7958)			
htd				-0.0056** (-2.3495)	0.0021*** (4.6246)	0.0014*** (4.6793)
ZSCORE		-0.0068*** (-3.4173)	-0.0022* (-1.9136)		-0.0068*** (-3.4484)	-0.0022* (-1.9139)
lev	-13.2629*** (-49.7632)	-0.1722*** (-2.8867)	-0.0887** (-2.2691)	-13.2883*** (-49.4168)	-0.1597*** (-2.6840)	-0.0814** (-2.0874)
ppe	-0.0157*** (-4.3388)	-0.0016** (-2.1492)	-0.0013*** (-2.6587)	-0.0156*** (-4.3210)	-0.0016** (-2.1826)	-0.0013*** (-2.6900)
roa	0.0203*** (4.4870)	-0.0019 (-1.6333)	-0.0026*** (-3.4857)	0.0195*** (4.2934)	-0.0016 (-1.3532)	-0.0024*** (-3.2251)
size	0.5667*** (10.2043)	0.0661*** (5.3760)	0.0311*** (3.7635)	0.6074*** (11.0928)	0.0484*** (4.0026)	0.0202** (2.4868)
dturn	-0.0505** (-2.4967)	-0.0881*** (-17.3595)	-0.0620*** (-18.3405)	-0.0465** (-2.2901)	-0.0904*** (-17.9847)	-0.0632*** (-18.8293)
tobinq	2.5848*** (46.4915)	0.0413*** (5.3422)	0.0108** (2.1193)	2.5950*** (46.7578)	0.0364*** (4.7441)	0.0079 (1.5648)
gdp	-0.7514*** (-7.1122)	-0.0256 (-0.9512)	0.0188 (1.0570)	-0.6481*** (-6.2883)	-0.0688** (-2.5683)	-0.0085 (-0.4791)
_cons	1.7865 (1.3419)	-1.4691*** (-5.2260)	-1.0782*** (-5.7059)	-0.1941 (-0.1668)	-0.6041** (-2.3713)	-0.5477*** (-3.2164)
N	24135	24135	24135	24135	24135	24135
R ²	0.7998	0.1813	0.1776	0.7998	0.1810	0.1776
F	664.9507	46.9886	52.9289	696.5087	45.7076	52.6387

5.5 Heterogeneity Analysis

Previous studies have shown that alleviating information asymmetry and improving corporate fundamentals are the mechanisms through which extreme temperatures influence the risk of stock price crashes. However, different geographical environments, firm characteristics, and corporate transparency levels have varying effects on the manifestation of information asymmetry and the improvement of corporate fundamentals. This paper conducts a heterogeneity analysis of the relationship between extreme temperatures and the risk of stock price crashes from three perspectives, which is geographical location, firm characteristics, and corporate governance.

5.5.1 Geographic Location

This study divides the sample into two groups (Eastern and Central-Western regions) based on the geographic location of the companies' headquarters. The results of the group-specific regression in Table 12 show that for the Eastern region group (columns 1, 3, 5, and 7), the coefficients for extreme temperatures are all significantly positive at the 1% level. In contrast, for the Central-Western region group (columns 2, 4, 6, and 8), although the coefficients for extreme temperatures are positive, they are not highly significant or are not significant at all. This suggests that, compared to the Midwest and West, listed companies in the East are more susceptible to the impact of extreme temperatures on the risk of stock price crashes. Possible reasons may be that the eastern region has a more concentrated industrial chain and higher-value assets, making it more sensitive to temperature fluctuations. Furthermore, its mature financial markets can interpret and price climate risk information more rapidly, whereas the information environment in the central and western regions is relatively lagging, and negative news may be concealed.

Table 8: Heterogeneity analysis based on geographic location

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Eastern	Midwest	Eastern	Midwest	Eastern	Midwest	Eastern	Midwest
	NCSKEW	NCSKEW	DUVOL	DUVOL	NCSKEW	NCSKEW	DUVOL	DUVOL
ltd	0.0047***	0.0024	0.0024***	0.0025**				
	(5.4093)	(1.5318)	(4.2208)	(2.3763)				
htd					0.0020***	0.0022**	0.0013***	0.0016**
					(3.9039)	(2.4575)	(3.8297)	(2.5826)
lev	-0.0424	-0.2111**	-0.0219	-0.1686**	-0.0283	-0.2014**	-0.0150	-0.1574**
	(-0.6662)	(-2.0614)	(-0.5117)	(-2.4696)	(-0.4450)	(-1.9819)	(-0.3519)	(-2.3230)
ppe	-0.0024***	0.0003	-0.0017***	-0.0004	-0.0024***	0.0003	-0.0017***	-0.0004
	(-2.6271)	(0.2586)	(-2.8946)	(-0.4724)	(-2.6040)	(0.2131)	(-2.8807)	(-0.5374)
roa	-0.0023*	-0.0010	-0.0027***	-0.0024	-0.0020	-0.0007	-0.0026***	-0.0021
	(-1.7233)	(-0.3956)	(-3.1532)	(-1.4959)	(-1.5049)	(-0.3063)	(-2.9626)	(-1.3507)
size	0.0627***	0.0599**	0.0285***	0.0316**	0.0426***	0.0498**	0.0174*	0.0223
	(4.3297)	(2.4697)	(2.9099)	(2.0097)	(2.9792)	(2.0966)	(1.8010)	(1.4509)
dturn	-0.0910***	-0.0788***	-0.0658***	-0.0516***	-0.0933***	-0.0805***	-0.0667***	-0.0536***
	(-15.5408)	(-7.6896)	(-16.7591)	(-7.7064)	(-16.0400)	(-8.0003)	(-17.0721)	(-8.0946)
tobinq	0.0243***	0.0212*	0.0058	0.0032	0.0177***	0.0199*	0.0026	0.0012
	(3.5848)	(1.7839)	(1.2452)	(0.3833)	(2.6650)	(1.6856)	(0.5567)	(0.1411)
gdp	0.0164	-0.1123**	0.0397*	-0.0305	-0.0377	-0.1406***	0.0088	-0.0538*
	(0.4933)	(-2.3113)	(1.7965)	(-0.9672)	(-1.1398)	(-2.8915)	(0.4015)	(-1.7071)
cons	-1.9214***	-0.3924	-1.2820***	-0.5349	-0.8574***	0.0865	-0.6974***	-0.0954
	(-5.4694)	(-0.8141)	(-5.4395)	(-1.6365)	(-2.6953)	(0.1976)	(-3.2895)	(-0.3237)
N	18066	6077	18066	6077	18066	6077	18066	6077
R ²	0.1836	0.1797	0.1804	0.1749	0.1829	0.1803	0.1803	0.1751
F	42.2208	10.9705	47.9432	12.3647	40.3536	11.2609	47.7602	12.3271

5.5.2 Information Disclosure

Poorly governed firms demonstrate significant deficiencies in information disclosure quality and investor relations management [50]. To examine the role of disclosure quality in the process by which extreme weather events influence the risk of stock price crashes, this paper categorizes the sample companies into different disclosure quality groups based on the disclosure assessment results published by the Shenzhen Stock Exchange and the Shanghai Stock Exchange. Specifically, companies with disclosure assessment results of "A" (Excellent) and "B" (Good) were classified into the high-quality disclosure group, while those with results of "C" (Pass) and "D" (Fail) were classified into the low-quality disclosure group. The higher the quality of a company's information disclosure, the higher its transparency.

The results in Table 13 show that columns (2), (4), (6), and (8) are significantly positive at the 1% level, indicating that the impact of extreme temperatures on the risk of stock price crashes is more pronounced in the high-quality information disclosure group. Companies with high-quality information disclosure have established a robust information environment, enabling them to convey climate risk-related information to the market in a timely and accurate manner. In this transparent information environment, investors can clearly identify the potential impact of extreme climate events on corporate fundamentals and generate a relatively

consistent market reaction, leading to a significant association between climate shocks and the risk of stock price crashes. Conversely, the information environment for companies with low-quality disclosure is relatively opaque, and climate risk information is drowned out by noise, making it difficult for investors to form consistent expectations. This, in turn, weakens the systemic impact of extreme climate events on the risk of stock price crashes.

Table 9: Heterogeneity analysis based on quality of information disclosure

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Low	High	Low	High	Low	High	Low	High
	NCSKEW	NCSKEW	DUVOL	DUVOL	NCSKEW	NCSKEW	DUVOL	DUVOL
ltd	0.0008	0.0056***	0.0017	0.0030***				
	(0.4856)	(6.0878)	(1.5108)	(5.0058)				
Htd					-0.0006	0.0029***	-0.0003	0.0018***
					(-0.6525)	(5.3601)	(-0.4547)	(4.8954)
lev	-0.1910	-0.0755	-0.1075	-0.0821*	-0.1835	-0.0665	-0.0949	-0.0773
	(-1.6304)	(-1.0735)	(-1.3974)	(-1.6974)	(-1.5810)	(-0.9451)	(-1.2449)	(-1.5976)
ppe	-0.0012	-0.0009	-0.0009	-0.0009	-0.0012	-0.0009	-0.0009	-0.0009
	(-0.7837)	(-0.9131)	(-0.8605)	(-1.3789)	(-0.7712)	(-0.9277)	(-0.8585)	(-1.3868)
roa	-0.0023	-0.0006	-0.0024	-0.0022**	-0.0023	-0.0002	-0.0023	-0.0020*
	(-1.0596)	(-0.3521)	(-1.6426)	(-2.1400)	(-1.0274)	(-0.0993)	(-1.5369)	(-1.9102)
size	0.0625**	0.0484***	0.0300	0.0237**	0.0589**	0.0225	0.0227	0.0089
	(2.1320)	(3.0623)	(1.5688)	(2.1950)	(2.0755)	(1.4397)	(1.2301)	(0.8394)
dturn	-0.0540***	-0.0958***	-0.0345***	-0.0682***	-0.0564***	-0.0987***	-0.0378***	-0.0696***
	(-4.8501)	(-15.3863)	(-4.5995)	(-16.4849)	(-5.2311)	(-15.9324)	(-5.1787)	(-16.8949)
tobinq	0.0029	0.0245***	-0.0102	0.0083*	0.0014	0.0175**	-0.0132	0.0046
	(0.2108)	(3.5186)	(-1.0577)	(1.6985)	(0.0997)	(2.5335)	(-1.3809)	(0.9492)
gdp	-0.1712***	-0.0095	-0.0948**	0.0327	-0.1753***	-0.0571	-0.1090***	0.0049
	(-2.8868)	(-0.2499)	(-2.4243)	(1.2933)	(-3.0603)	(-1.5007)	(-2.8937)	(0.1937)
cons	0.2750	-1.3185***	0.2008	-1.0818***	0.4503	-0.2346	0.5778	-0.4688**
	(0.3749)	(-3.5747)	(0.4033)	(-4.3563)	(0.7158)	(-0.6739)	(1.3665)	(-1.9985)
N	6076	17509	6076	17509	6076	17509	6076	17509
R ²	0.2482	0.2203	0.2470	0.2133	0.2482	0.2199	0.2467	0.2133
F	6.1086	41.1591	7.5167	44.8556	6.1104	39.3255	7.3033	44.3973

5.5.3 Corporate property rights

Under China's distinctive institutional background, there are differences between state-owned firms and private firms in terms of resource acquisition capabilities, government support, and credibility^[51-53]. The sample was divided into two groups based on corporate property rights—the state-owned enterprises and non-state-owned enterprises. The results in Table 14 indicate that non-state-owned enterprises exhibit a more significant impact compared to state-owned enterprises. A possible reason is that state-owned enterprises possess inherent advantages in the credit market, enabling them to obtain bank loans and other financing support at lower costs. Following a climate shock, state-owned enterprises can readily access reconstruction funds to maintain normal operations. In contrast, non-state-owned enterprises face financing constraints. Therefore, a climate shock may exacerbate their financing difficulties and amplify operational risks.

Table 10: Heterogeneity analysis based on corporate property rights

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	nonstate-owned	State-owned	nonstate-owned	State-owned	nonstate-owned	State-owned	nonstate-owned	State-owned
	NCSKEW	NCSKEW	DUVOL	DUVOL	NCSKEW	NCSKEW	DUVOL	DUVOL
ltd	0.0051***	0.0029**	0.0029***	0.0019**				
	(4.9627)	(2.5443)	(4.2465)	(2.4824)				
htd					0.0025***	0.0012*	0.0016***	0.0009*
					(4.2382)	(1.6926)	(3.9304)	(1.9578)
lev	-0.0712	-0.0866	-0.0541	-0.0629	-0.0684	-0.0624	-0.0527	-0.0476
	(-1.0185)	(-0.8961)	(-1.1419)	(-0.9897)	(-0.9783)	(-0.6510)	(-1.1117)	(-0.7553)
ppe	-0.0005	-0.0033***	-0.0008	-0.0023***	-0.0005	-0.0033***	-0.0009	-0.0023***
	(-0.4813)	(-2.9791)	(-1.2048)	(-3.1409)	(-0.5278)	(-2.9668)	(-1.2492)	(-3.1223)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	nonstate-owned	State-owned	nonstate-owned	State-owned	nonstate-owned	State-owned	nonstate-owned	State-owned
	NCSKEW	NCSKEW	DUVOL	DUVOL	NCSKEW	NCSKEW	DUVOL	DUVOL
roa	-0.0018	-0.0043*	-0.0025***	-0.0037**	-0.0015	-0.0038*	-0.0024**	-0.0034**
	(-1.2769)	(-1.9074)	(-2.7226)	(-2.4768)	(-1.0850)	(-1.6765)	(-2.5492)	(-2.2533)
size	0.0951***	0.0350*	0.0432***	0.0172	0.0731***	0.0220	0.0302***	0.0087
	(5.7369)	(1.7148)	(3.8331)	(1.2765)	(4.4815)	(1.0967)	(2.7225)	(0.6560)
dturn	-0.0958***	-0.0721***	-0.0706***	-0.0461***	-0.0983***	-0.0747***	-0.0719***	-0.0477***
	(-14.1501)	(-8.8213)	(-15.6758)	(-8.3510)	(-14.6151)	(-9.2744)	(-16.0439)	(-8.7216)
tobinq	0.0218***	0.0250**	0.0050	0.0007	0.0152**	0.0218*	0.0013	-0.0012
	(3.0663)	(2.2249)	(1.0041)	(0.0909)	(2.1802)	(1.9537)	(0.2611)	(-0.1529)
gdp	-0.0731*	0.0141	0.0006	0.0379	-0.1249***	-0.0133	-0.0303	0.0189
	(-1.8020)	(0.3580)	(0.0226)	(1.4789)	(-3.0790)	(-0.3393)	(-1.1219)	(0.7404)
cons	-1.5467***	-1.3055***	-1.1060***	-1.0374***	-0.4933	-0.7024*	-0.4879*	-0.6372**
	(-3.8004)	(-2.9663)	(-3.9911)	(-3.5361)	(-1.3074)	(-1.7631)	(-1.9070)	(-2.4271)
N	14176	9362	14176	9362	14176	9362	14176	9362
R ²	0.2083	0.1628	0.1998	0.1611	0.2078	0.1625	0.1997	0.1609
F	36.2897	14.2473	43.0372	15.4305	35.0697	13.7818	42.6733	15.1336

6. Conclusion and Implications

This study uses non-financial A-share listed companies on the Shanghai and Shenzhen stock exchanges from 2010 to 2023 as the research sample to empirically examine the impact of extreme temperatures on stock price crash risk and its underlying mechanisms. The findings reveal that extreme temperatures can significantly exacerbate a company's stock price crash risk, with this effect being more pronounced in companies located in the eastern region, those with high disclosure levels, and non-state-owned enterprises. Furthermore, this conclusion remains valid after a series of endogeneity and robustness tests. The results of the mechanism analysis indicate that extreme temperatures exacerbate the risk of stock price crashes by increasing the degree of information asymmetry and deteriorating corporate fundamentals.

Based on the above findings, this paper offers the following policy implications. For regulatory authorities, it is essential to further strengthen disclosure quality requirements for listed companies, particularly regarding environmental risk-related information, to enhance the market's efficiency in pricing climate risks. Financial regulators may also collaborate with meteorological agencies to develop high-resolution, enterprise-level climate risk databases and early-warning platforms, helping small and medium-sized enterprises identify and quantify their climate risks at lower costs.

For companies, based on their level of climate risk exposure, they should prudently assess and maintain an appropriate level of precautionary cash reserves to address liquidity shocks caused by extreme weather events. Companies need to actively seek stable capital and recognize the "stabilizing" role of institutional investors. By improving corporate governance and strengthening communication, proactively attracting and retaining institutional investors with a long-term perspective and oversight capabilities can help stabilize stock prices during market turbulence and cushion external shocks.

Investors should place high importance on the quality of disclosure in investment portfolios and incorporate quantitative analyses of companies' climate risk exposure and adaptive capacity into valuation models and credit ratings to better manage climate risk exposure. They should also pay particular attention to companies' assets in eastern regions, their non-state-owned status, and the adequacy of their disclosures. Institutional investors should not merely passively avoid risks; rather, they should leverage their shareholder rights to urge companies to improve climate risk management and disclosure through shareholder proposals, voting, and dialogue with management.

References

- [1] The IPCC Sixth Assessment Report on Climate Change Impacts [J]. Population and Development Review, 2022,48(2):629-633.

- [2] Jin L, Myers S C. R2 around the world: New theory and new tests [J]. *Journal of Financial Economics*, 2006,79(2):257-292.
- [3] Kim J B, Zhang L. Financial Reporting Opacity and Expected Crash Risk: Evidence from Implied Volatility Smirks [J]. *Contemporary Accounting Research*, 2014,31(3):851-875.
- [4] Hutton A P, Marcus A J, Tehranian H. Opaque financial reports, R2, and crash risk [J]. *Journal of Financial Economics*, 2009,94(1):67-86.
- [5] Francesca D, Barbara A, Matthias K, et al. Climate actions and macro-financial stability: The role of central banks [J]. *Journal of Environmental Economics and Management*, 2021,110.
- [6] Xu G, You W, Boqiang L. Assessing dynamic China's energy security: Based on functional data analysis [J]. *Energy*, 2020(prepublish):119324.
- [7] BOLTON P, KACPERCZYK M. Global Pricing of Carbon-Transition Risk [J]. *The Journal of Finance*, 2023,78(6):3677-3754.
- [8] Alharbi S S, Tabash M I, Farooq U, et al. How does climate policy uncertainty determine green innovation adoption? New perspectives from the BRICS [J]. *The Journal of Economic Asymmetries*, 2025,31:e00411.
- [9] Raza S A, Khan K A, Benkraiem R, et al. The importance of climate policy uncertainty in forecasting the green, clean and sustainable financial markets volatility [J]. *International Review of Financial Analysis*, 2024,91:102984.
- [10] Zhang D. Subsidy expiration and greenwashing decision: Is there a role of bankruptcy risk? [J]. *Energy Economics*, 2023,118:106530.
- [11] CHEN Guojin D S Z X. Climate Policy, Bank Risk, and Macro-Prudential Regulation [J]. *Journal of Financial Research*, 2023,519(9):38-57.
- [12] Easaw J, Grimme C. The relationship between aggregate uncertainty and firm-level uncertainty [J]. *Oxford Economic Papers*, 2024,76(4):1108-1127.
- [13] An Z, Chen C, Naiker V, et al. Does media coverage deter firms from withholding bad news? Evidence from stock price crash risk [J]. *Journal of Corporate Finance*, 2020,64.
- [14] Li Y, Zeng Y. The impact of top executive gender on asset prices: Evidence from stock price crash risk [J]. *Journal of Corporate Finance*, 2019,58:528-550.
- [15] Callen J L, Fang X. Institutional investor stability and crash risk: Monitoring versus short-termism? [J]. *Journal of Banking and Finance*, 2013,37(8):3047-3063.
- [16] Li X, Wang S S, Wang X. Trust and stock price crash risk: Evidence from China [J]. *Journal of Banking and Finance*, 2017,76:74-91.
- [17] Hou C, Liu H. Institutional cross-ownership and stock price crash risk [J]. *Research in International Business and Finance*, 2023,65:101906.
- [18] Zhou J, Zhu Z, Wang H, et al. Data Asset Disclosure and Stock Price Crash Risk: A Double Machine Learning Study of Chinese A Share Firms: *International Journal of Financial Studies* [Z]. 2025: 13, 229.
- [19] Kothari S P, Shu S, Wysocki P D. Do Managers Withhold Bad News? [J]. *Journal of Accounting Research*, 2009,47(1):241-276.
- [20] Yuan D, Shao J. Rookie Chinese directors and stock price crash risk [J]. *Management Decision*, 2025.
- [21] Liu L, Li Y. Enterprise financialization and stock price crash risk [J]. *Finance Research Letters*, 2024,60:104878.
- [22] Zhang X, Wang B, Liu G. Readability of financial reports and stock price crash risk [J]. *Finance Research Letters*, 2025,86:108489.
- [23] Fu F, Fang J, Yang M, et al. Institutional investor horizons and stock price crash risk [J]. *Research in International Business and Finance*, 2024,72:102509.

- [24] Pu G, Xie Y, Wang K. Board faultlines and risk-taking [J]. *Finance Research Letters*, 2023,51:103404.
- [25] DAI Bingbin, YUE Heng. Monetary Policy, Liquidity Shortage and Stock Price Crash Risk. *Journal of Financial Research*, 2015, 421(7): 135-151.
- [26] Shao H, Wang Y, Wang Y, et al. Green credit policy and stock price crash risk of heavily polluting enterprises: Evidence from China [J]. *Economic Analysis and Policy*, 2022,75:271-287.
- [27] Ye Y, Zeng L, Tao Y, et al. Tax authority monitoring and corporate information disclosure quality in China [J]. *International Review of Financial Analysis*, 2023,90:102872.
- [28] Stroebel J, Wurgler J. What do you think about climate finance? [J]. *Journal of Financial Economics*, 2021,142(2):487-498.
- [29] Naumann G, Cammalleri C, Mentaschi L, et al. Increased economic drought impacts in Europe with anthropogenic warming [J]. *Nature Climate Change*, 2021,11(6):485-491.
- [30] Massa M, Zhang L. The Spillover Effects of Hurricane Katrina on Corporate Bonds and the Choice between Bank and Bond Financing [J]. *Journal of Financial and Quantitative Analysis*, 2020,56(3):1-57.
- [31] Choi D, Gao Z, Jiang W. Attention to Global Warming [J]. *The Review of Financial Studies*, 2020,33(3):1112-1145.
- [32] D. H T, Ying X. Panic Selling When Disaster Strikes: Evidence in the Bond and Stock Markets [J]. *Management Science*, 2021.
- [33] Rye C J, Boyd J A, Mitchell A. Normative approach to risk management for insurers [J]. *Nature Climate Change*, 2021,11(6):460-463.
- [34] M. A J, T. N D, Ariel O B. Temperature shocks and industry earnings news [J]. *Journal of Financial Economics*, 2023,150(1):1-45.
- [35] Ojea Ferreiro J, Reboredo J, Ugolini A. The impact of climate transition risks on financial stability. A systemic risk approach [M]. 2022.
- [36] Huang H H, Kerstein J, Wang C. The impact of climate risk on firm performance and financing choices: An international comparison [J]. *Journal of International Business Studies*, 2018,49(5):633-656.
- [37] LV Yongbin L Z G Y. Marching Against the Wind: Typhoon Disasters and Banks' Risky Behaviors [J]. *Journal of Financial Research*, 2024,523(1):19-37.
- [38] Huang R, Xie X, Zhou H. 'Isomorphic' behavior of corporate greenwashing [J]. *Chinese Journal of Population, Resources and Environment*, 2022,20:29-39.
- [39] Al Mamun M, Balachandran B, Duong H N. Powerful CEOs and stock price crash risk [J]. *Journal of Corporate Finance*, 2020,62:101582.
- [40] Wang, H.C., Cao, F. and Ye, K.T. Monitoring or Tunneling? The Proportion of the Proportion Held by the Big Shareholders and the Risk of the Crash of the Stock Price [J]. *Management World*, 2015,02:45-57.
- [41] Kim J B, Lee E, Zhu Z J. Do Firm-Specific Stock Price Crashes Lead to a Stimulation or Distortion of Market Information Efficiency? [J]. *Contemporary Accounting Research*, 2022,39(3):2175-2211.
- [42] Chang X, Chen Y, Zolotoy L. Stock Liquidity and Stock Price Crash Risk [J]. *The Journal of Financial and Quantitative Analysis*, 2017,52(4):1605-1637.
- [43] Habib A, Hasan M M, Jiang H. Stock price crash risk: review of the empirical literature [J]. *Accounting & Finance*, 2018,58:211-251.
- [44] Ziqi Z, Zhi S, Ke W, et al. Corporate environmental information disclosure and stock price crash risk: Evidence from Chinese listed heavily polluting companies [J]. *Energy Economics*, 2022,112.
- [45] He G, Bai L, Ren H M. Analyst coverage and future stock price crash risk [J]. *Journal of Applied Accounting Research*, 2019,20(1):63-77.

- [46] He G, Ren H. Are financially constrained firms susceptible to a stock price crash? [J]. The European Journal of Finance, 2022,29:1-26.
- [47] Yu H, Liang C, Liu Z, et al. News-based ESG sentiment and stock price crash risk [J]. International Review of Financial Analysis, 2023,88:102646.
- [48] Rao S, Koirala S, Thapa C, et al. When rain matters! Investments and value relevance [J]. Journal of Corporate Finance, 2022,73:101827.
- [49] Ni X, Si Y, Zhang B. Climate vulnerability and stock price crash risk worldwide [J]. Journal of International Business Studies, 2025:1-24.
- [50] Yin Z, Li X, Si D, et al. China stock market liberalization and company ESG performance: The mediating effect of investor attention [J]. Economic Analysis and Policy, 2023,80:1396-1414.
- [51] Kong D, Kong G, Liu S, et al. Does competition cause government decentralization? The case of state-owned enterprises [J]. Journal of Comparative Economics, 2022,50(4):1103-1122.
- [52] Li X L, Si D K. Does financial market liberalization promote corporate radical innovation? Evidence from China [J]. International Review of Financial Analysis, 2024,95:103500.
- [53] Si D K, Liu G. How does capital market liberalization shape corporate shadow banking? Evidence from China [J]. China Economic Review, 2025,93:102484.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

Integrating ESG into Factor Pricing Models and Actuarial Risk Estimation: Implications for Venture Capital

Linlang Song*

Nanjing University of Information Science & Technology, Nanjing, China

**Corresponding author: Linlang Song.*

Abstract

This paper reviews how Environmental, Social and Governance (ESG) factors can be incorporated into factor pricing models and actuarial risk estimation methods, and evaluates their implications for venture capital (VC) investment. Traditional asset pricing models, including single-factor and multi-factor models, provide a framework for assessing systematic risk, while actuarial techniques such as Value at Risk (VaR), Conditional Value at Risk (CVaR), and Extreme Value Theory (EVT) are effective in measuring tail risk. Existing literature suggests that ESG performance influences both expected returns and downside risk, with firms exhibiting stronger ESG profiles generally showing lower tail risk and greater resilience during periods of financial stress. In the context of venture capital, ESG serves as a signal of management quality, governance capability, and long-term growth potential, helping investors mitigate information asymmetry and improve investment selection. This review identifies major research findings, ongoing debates, and current gaps, particularly the lack of an integrated framework combining factor pricing and actuarial methods for ESG analysis in private markets. The paper concludes by proposing future research directions for building quantitative ESG assessment tools tailored to venture capital.

Keywords

environmental, social and governance (ESG), factor pricing model, actuarial risk estimation, venture capital, tail risk

1. Introduction

In recent years, Environmental, Social and Governance (ESG) investing has grown rapidly worldwide. According to the Global Sustainable Investment Alliance (GSIA), global sustainable investment assets have exceeded USD 30 trillion, and ESG factors have become an increasingly important component of asset allocation and risk management in capital markets. In this context, venture capital (VC) firms have gradually incorporated ESG considerations into their investment decision-making processes, as start-up companies are highly dependent on external financing in the early stages, and ESG performance is widely regarded as an indicator of long-term growth potential, governance quality, and potential risk exposure [1,2]. Despite the rapid development of ESG investing in practice, traditional quantitative financial models, such as factor pricing models and actuarial risk estimation methods, have not yet been systematically applied to ESG risk analysis.

This methodological gap makes it difficult to accurately capture and measure ESG-related risks, particularly tail risk and systematic risk, thereby limiting the depth of ESG applications in asset pricing and venture capital decision-making. Therefore, it is necessary to develop a systematic modeling framework for ESG risk by integrating financial econometrics and actuarial risk theory [1]. Existing research on factor pricing models, actuarial risk estimation, and ESG finance remains highly fragmented. Factor pricing models provide a fundamental framework for asset pricing and risk decomposition, but they require structural extensions to incorporate ESG factors and account for their nontraditional risk characteristics [3-5]. Meanwhile, actuarial risk estimation methods, including Value at Risk (VaR), Conditional Value at Risk (CVaR), and Extreme Value Theory (EVT), are particularly effective in analyzing low-probability, high-impact events and are therefore well suited to capturing the tail-risk characteristics of ESG-related events. Integrating actuarial methods into ESG risk measurement can improve the identification of extreme risk exposures arising from environmental and social shocks. Against this background, this paper examines the role of ESG in quantitative finance and venture capital, focusing on three key questions: how factor pricing models can incorporate ESG factors to capture systematic risk; how actuarial risk estimation methods can be used to quantify ESG-related tail risk; and how ESG integration influences venture capital performance and the mechanisms through which it operates. Building on these questions, this study integrates financial econometrics, actuarial science, and sustainable finance theory to develop a conceptual framework for quantitative ESG analysis in venture capital and to address the current lack of interdisciplinary integration in the literature. The remainder of this paper is organized as follows. Section 2 introduces the theoretical foundations, including factor pricing theory, actuarial risk estimation methods, and the theoretical perspectives that explain ESG as a risk driver. Section 3 provides a systematic literature review covering ESG in asset pricing models, ESG-related risk measurement, and ESG in venture capital. Section 4 synthesizes the literature and discusses the main areas of consensus, debate, and research gaps. Section 5 concludes and outlines directions for future research.

2. Theoretical Foundations

2.1 Factor Pricing Models

Factor pricing models are central tools in modern financial economics. The single-index model developed by Sharpe decomposes asset returns into systematic market risk and idiosyncratic risk [3]. Subsequently, multifactor models such as the Fama-French three-factor model and the Carhart four-factor model introduced additional factors including size, value, and momentum to explain cross-sectional differences in returns [4,5]. These models distinguish between systematic risk and idiosyncratic risk, where the former is driven by common macroeconomic factors and the latter can be diversified away. This framework provides a natural foundation for ESG analysis. If ESG characteristics influence the return patterns of a broad set of firms, ESG may be interpreted as a systematic factor and incorporated into traditional pricing models [1].

2.2 Actuarial Risk Estimation

The core objective of actuarial risk estimation is to quantify extreme losses. Value at Risk (VaR) measures the maximum expected loss at a given confidence level, while Conditional Value at Risk (CVaR), also known as Expected Shortfall, estimates the average loss beyond the VaR threshold. Because actuarial science focuses on low-probability, high-severity events, it is particularly well suited to analyzing ESG risks such as environmental disasters, labor controversies, and governance scandals. CVaR and Extreme Value Theory (EVT) provide effective tools for assessing tail risks associated with ESG events.

2.3 ESG as a Risk Driver

From a theoretical perspective, ESG can be interpreted as a risk driver through several complementary lenses. Stakeholder theory argues that firms must create value not only for shareholders but also for employees, customers, suppliers, and communities, and that effective stakeholder management reduces operational and reputational risk [1]. Signaling theory suggests that ESG performance functions as a signal of management quality and long-term orientation, allowing firms to communicate stronger governance and risk-control capabilities to investors [1,2]. Legitimacy theory further posits that firms require societal approval to maintain operations, and ESG practices help firms align with social norms, thereby reducing regulatory, reputational, and consumer-related risks [1].

3. Literature Review

3.1 How ESG Enters Asset Pricing

A large body of empirical evidence indicates that ESG performance is positively associated with corporate financial performance. Friede et al. conducted a meta-analysis covering more than 2,000 empirical studies and found that approximately 90% reported a non-negative relationship between ESG and corporate financial performance [6]. Gillan et al. further argue that ESG affects firm value through improved cash flows, lower financing costs, reduced regulatory and reputational risks, and stronger investor demand [7]. In addition, Bolton and Kacperczyk find that firms with higher carbon emissions tend to earn higher expected returns, suggesting that climate-related risk is priced in capital markets [8]. Overall, ESG enters asset pricing through both firm fundamental channels and investor preference channels.

3.2 How ESG Affects Risk Distribution

Beyond expected returns, ESG also influences the distribution of returns, particularly downside and tail risk. Broadstock et al. show that firms with stronger ESG performance exhibit lower downside risk and greater resilience during financial stress periods [9]. Ding et al. further find that firms with higher ESG scores experienced smaller stock price declines during the COVID-19 pandemic and more stable operating performance [10]. These results indicate that ESG reduces left-tail risk and enhances financial resilience. Tail risk analysis further suggests that firms with strong ESG performance are less exposed to extreme negative outcomes due to reduced environmental controversies, governance failures, and social conflicts. In addition, stakeholder relationships formed through ESG practices can act as a stabilizing mechanism during market stress [1,9]. In methodological terms, Extreme Value Theory (EVT) is widely used to model rare but severe ESG-related losses, as it better captures fat-tailed distributions than normality-based models.

3.3 ESG in Venture Capital

In venture capital markets, ESG is particularly relevant due to high uncertainty and information asymmetry. Startups typically face higher failure rates and lack stable financial histories, making traditional financial indicators insufficient for evaluation. Kaplan and Strömberg show that venture capital investing relies heavily on governance structures, control rights, and founder quality to mitigate information asymmetry [11]. Gompers et al. further demonstrate that venture capitalists systematically evaluate management quality and governance characteristics, which significantly affect investment decisions and exit outcomes [12]. Within this context, ESG can be interpreted as an extension of soft information indicators. It provides a multidimensional assessment of governance quality, social responsibility, and environmental sustainability. Dimson et al. show that ESG engagement improves governance and operational performance [13], while Krüger finds that ESG-related information significantly influences investor expectations and firm valuation [14]. Overall, ESG contributes to venture capital through signaling effects, growth identification, and downside risk reduction.

4. Discussion

4.1 Consensus Across Studies

Existing research generally agrees that ESG factors exhibit systematic risk characteristics and influence asset pricing at the market level rather than only at the firm-specific level [1]. There is also a broad consensus that higher ESG performance is associated with lower tail risk and stronger resilience during extreme market conditions [9,10]. Furthermore, most studies indicate that ESG integration improves portfolio risk-adjusted performance by reducing downside risk, even though its effect on abnormal returns remains debated [6,7].

4.2 Contradictions and Debates

However, several key debates remain unresolved. First, the existence and direction of the ESG premium are still inconclusive, with empirical results varying across markets and methodologies [6,8]. Second, there is no consensus on whether ESG represents an independent risk factor or merely reflects investor preferences, which affects its theoretical stability. Third, ESG rating systems exhibit relatively low cross-agency

consistency, with correlations typically between 0.5 and 0.7, which limits the robustness of empirical factor construction.

4.3 Mechanisms and Research Trends

The literature shows three major shifts in research focus. First, ESG research has moved from exclusion-based screening strategies to systematic factor integration into formal asset pricing models [1]. Second, attention has shifted from mean–variance frameworks to tail risk analysis, emphasizing extreme event behavior [9,10]. Third, research has expanded from public markets to private markets, particularly venture capital and private equity, where ESG is increasingly studied as part of investment decision-making [11,12].

4.4 Research Gaps and Limitations

Despite progress, several gaps remain. There is still no unified framework integrating factor pricing models with actuarial risk methods. ESG data availability in venture capital remains limited due to the private nature of firms and the lack of standardized disclosure. Existing studies are mostly cross-sectional and fail to capture long-term dynamics across the full investment lifecycle. Finally, ESG mechanisms in high-growth startups remain underexplored, as most theories are based on mature firm environments.

4.5 Reflections from Project Experience

Empirical practice shows that multi-factor models can incorporate ESG factors, but industry heterogeneity poses challenges in weighting ESG dimensions. Manufacturing industries tend to emphasize environmental factors, while service industries focus more on social dimensions such as data privacy and labor rights. Additionally, Tail Value at Risk (TVaR) is particularly useful for capturing extreme ESG-related risks, although its practical application in venture capital remains constrained by data limitations and model complexity.

5. Conclusion

This paper reviews the integration of ESG into factor pricing models and actuarial risk estimation. The literature evolves from single-factor to multi-factor models, from mean–variance analysis to tail risk focus, and from public to private markets. While ESG is widely associated with lower tail risk, its effect on returns remains uncertain. This study proposes an integrated framework combining financial econometrics, actuarial risk methods, and ESG theory, and highlights its relevance for venture capital risk assessment. It also summarizes key consensus areas and unresolved debates in the literature. Future research should focus on integrating factor models with tail-risk frameworks, conducting longitudinal VC lifecycle studies, developing industry-specific ESG factors, and applying machine learning techniques to improve ESG risk estimation under data constraints.

References

- [1] Bani-Khaled, S., Azevedo, G., & Oliveira, J. (2025). Environmental, social, and governance (ESG) factors and firm value: A systematic literature review of theories and empirical evidence. *AMS Review*, 15, 228–260.
- [2] Aldukkah, A. (2023). The impact of management quality on the ESG risks [Master’s thesis]. Hamad Bin Khalifa University.
- [3] Sharpe, W. F. (1963). A simplified model for portfolio analysis. *Management Science*, 9(2), 277–293.
- [4] Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, 33(1), 3–56.
- [5] Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, 52(1), 57–82.
- [6] Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: Aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210–233.

- [7] Gillan, S. L., Koch, A., & Starks, L. T. (2021). Firms and social responsibility: A review of ESG and CSR research in corporate finance. *Journal of Corporate Finance*, 66, 101889.
- [8] Bolton, P., & Kacperczyk, M. (2021). Do investors care about carbon risk? *Journal of Financial Economics*, 142(2), 517–549.
- [9] Ding, X., et al. (2021). Corporate immunity to the COVID-19 pandemic. *Journal of Financial Economics*.
- [10] Broadstock, D. C., et al. (2021). The role of ESG performance during times of financial crisis. *Journal of Banking & Finance*.
- [11] Kaplan, S. N., & Strömberg, P. (2003). Financial contracting theory meets the real world: An empirical analysis of venture capital contracts. *Review of Economic Studies*, 70(2), 281–315.
- [12] Gompers, P. A., Gornall, W., Kaplan, S. N., & Strebulaev, I. A. (2016). How do venture capitalists make decisions? *Journal of Financial Economics*, 121(1), 1–23.
- [13] Dimson, E., Karakaş, O., & Li, X. (2020). Coordinated engagements. *Journal of Financial Economics*, 138(2), 576–595.
- [14] Krüger, P. (2015). Corporate goodness and shareholder wealth. *Journal of Financial Economics*, 115(2), 304–329.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights

Copyright for this article is retained by the author (s), with first publication rights granted to the journal. This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).